Transitive, Anti-Symmetric Relational Attributes in Structural Description Matching with Applications to Radar Target Identification

O.S. Sands and F.D. Garber

The Ohio State University

ElectroScience Laboratory

Department of Electrical Engineering
Columbus, Ohio 43212

Technical Report 723090-1
Grant N00014-90-J-1200
October 1990

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

50272-101

REPORT DOCUMENTATION PAGE	1. REPORT NO.	2.	3. Recipient's Accession No.

4. Title and Subtitle
Transitive, Anti-Symmetric Relational Attributes in Structural Description Matching with Applications to Radar Target Identification

5. Report Date
October 1990

6.

7. Author(s)
O.S. Sands and F.D. Garber

8. Performing Org. Rept. No.
723090-1

9. Performing Organization Name and Address
The Ohio State University
ElectroScience Laboratory
1320 Kinnear Road
Columbus, OH 43212

10. Project/Task/Work Unit No.

11. Contract(C) or Grant(G) No.
(C)
(G) N00014-90-J-1200

12. Sponsoring Organization Name and Address
Office of Naval Research
800 North Quincy Street, Code 1512A:EAM
Arlington, VA 22217-5000

13. Report Type/Period Covered
Technical Report

14.

15. Supplementary Notes

16. Abstract (Limit: 200 words)

A structural pattern recognition system for a radar target identification system is described. Segmentation of a radar measurement series is accomplished with a Prony-based estimation procedure. Matching of structural descriptions is developed for structural descriptions with specialized relational attributes and for which an inter-node-set distance function is available. The reformulated matching theory is applied to the radar target identification problem.

The radar measurement process is described in detail. A parametric model for the measurement series is given. An estimation process for the parametric model is briefly described and the energy localization properties of the estimation process are made explicit. The effects of noise on a parametric decomposition are discussed.

An alternative framework for the matching of structural descriptions is developed. This development includes a number of qualifications applied to the standard structural description framework. A required inter-node-set distance function is used to define an inter-structural-description distance. Qualifications applied to the relational portion of the structural description formalism lead to the use of registration parameters for resolution of relational discrepancies.

The newly developed matching framework is applied to the target identification problem. A classifier based on the described matching strategy is shown to yield performance rivalling a statistical classifier under the assumption of known noise environment and aspect and superior performance in the presence of extraneous or missing responses.

17. Document Analysis a. Descriptors

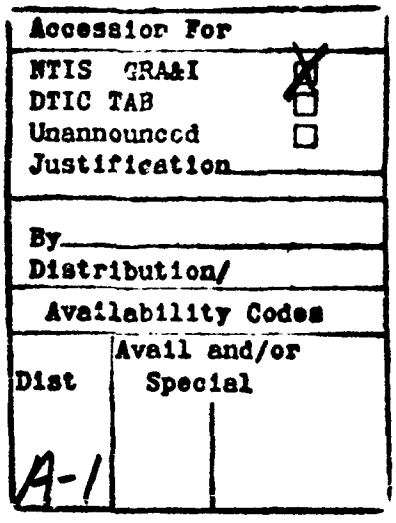
TARGET IDENTIFICATION
RELATIONAL GRAPH MATCHING
PRONY MODEL
PATTERN RECOGNITION
PARAMETRIC SPECTRAL ESTIMATION
DISTANCE FUNCTIONS

b. Identifiers/Open-Ended Terms

c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
A. Approved for public release; Distribution is unlimited.	Unclassified	138
	20. Security Class (This Page) Unclassified	22. Price

(See ANSI-Z39.18) *See Instructions on Reverse* OPTIONAL FORM 272 (4-77)
Department of Commerce



Contents

List of Figures

List of Tables

List of Symbols

Symbol	Description	Page
l	Number of cycles per sub-pulse	13
N	Number of sub-pulses in a pulse number of measurements	13
τ	Pulse length	13
f_o	Initial frequency	14
Δf	Frequency spacing	14
f_k	Frequency corresponding to k^{th} measurement	14
y_k	k^{th} measurement	14
Υ	Measurement vector	14
R	Unambiguous range	14
c	Speed of light	14
δr	Range cell size	17

p_i	Pole	18
d_i	Residue	18
M	Model order	18
Λ	A parametric decomposition	19
v	Vector form of a node	19
x_i	Mode range location	20
ρ	Pole modulus	26
β	Proportion of the unambiguous range interval	26
E	Proportion of the total energy in a mode	26
γ	The half-power rangewidth	26
$\mathbf{D}$	A structural description	40
$\mathbf{P}$	A node set	40
$\mathbf{R}$	A set of relations	40
T	A set of node attributes	40
V_i	The set of *value* sets	40
a_i	A node attribute	40
s	An element of the node set	40
v_i	An element of the set of node attribute value sets	40

r_k	A relation; an element of $\mathbf{R}$	40
NR_k	The *name* of relation r_k	40
R_k	Set of N-tuples which define relation r_k	40
f_{r_k}	Relational attribute function	40
$d_s(A,B)(\vec{r})$	The distance between node set A and node set B evaluated at registration parameter vector $\vec{r}$	48
$\vec{r}$	A vector of registration parameters	48
$\mathcal{A}$	A partition of the node set of $\mathbf{D}_A$	48
$\mathcal{B}$	A partition of the node set of $\mathbf{D}_B$	48
$\mathcal{M}$	A mapping from the elements of $\mathcal{A}$ to the elements of $\mathcal{B}$	48
$d_p(\mathcal{A},\mathcal{B})$	The distance between the partitions $\mathcal{A}$ and $\mathcal{B}$	49
$\mathcal{M}(\mathcal{A})$	The image of $\mathcal{A}$	49
$\mathcal{M}^{-1}(\mathcal{B})$	The pre-image of $\mathcal{B}$	49
$d\,(\mathbf{D}_A,\mathbf{D}_B)$	The distance between structural descriptions $\mathbf{D}_A$ and $\mathbf{D}_B$	49
$\mathbf{Part(U)}$	The set of partitions of the set $\mathbf{U}$	49
$\mathcal{P}$	A partition of the combined sets $\mathbf{P}_A \cup \mathbf{P}_B$	53
π	An element of the partition $\mathcal{P}$	54
π_A	The projection of π onto the node set $\mathbf{P}_A$	54
π_B	The projection of π onto the node set $\mathbf{P}_B$	54

$\mathbf{L}(\mathcal{P})$	The number of links implied by the partition $\mathcal{P}$	60
α	Link normalization factor	60
ϕ	The nil-map propensity	60
$\boldsymbol{\mathcal{P}}$	Matrix of vector sums of vector form of modes, implied by partition $\mathcal{P}$	108
$\delta_{\mathcal{P}}$	Generalized d-matrix corresponding to the partition $\mathcal{P}$	108

Chapter 1

INTRODUCTION

Non-cooperative radar target identification has been a topic of interest for some time [1]. This topic has been primarily motivated by the ability of a radar system to detect objects at distances beyond the detection range of other sensors.

Statistical methods of pattern recognition have been successfully applied to the target identification problem [2,3,4]. An underlying assumption of statistical methods is that the pattern generation process can be modeled as a vector valued random process [5]. Such methods produce "optimum" decision rules in cases when the assumed probabilistic modeling applies exactly [6].

In contrast to the statistical approach to pattern recognition, there exists a set of alternative methods to pattern recognition which are collectively referred to as *structural* methods of pattern recognition [7,8]. The defining property for a structural method is that the modeling of the pattern generation process is accomplished with a *symbolic* procedure. That is, pattern representations presented to a classifier consist of a number of discrete "en-

tities" or "primitives" with accompanying information regarding the type and inter-relationships among the constituent primitives. In the structural approach, the symbolic pattern representations are designed to embody information pertaining to object "structure" or "form" as opposed to specifically describing all of the details of the observed manifestation [7,9]; thus the term "structural".

This investigation is concerned with the formulation and application of a specific structural method of pattern recognition to the identification of airborne targets using a radar sensor. The purpose of this application is twofold. Of primary concern is the demonstration of the benefits of a structural approach to target identification. In addition to the benefits which accompany the use of a structural approach, the structural classifier will be shown to exhibit superior robustness compared to a statistical approach with respect to unmodeled target responses.

Generation of radar targets can be modeled fairly well with a probabilistic model. A classifier can therefore be defined which provides statistically optimal performance. A second goal of this study is to provide an example of how a structural approach can be applied to a problem such as radar target identification with little or no compromise in classifier performance, thereby establishing the applicability of the structural approach to such problems.

1.1 Structural Methods of Pattern Recognition for Sensor Data

The use of structural methods for recognition of sensor data from unknown objects has been proposed previously [7,10,11,12]. Structural methods have classically been applied in circumstances in which a probabilistic modeling of the pattern generation process is impractical. In these cases, extraction of symbolic pattern representations from sensor data (the *segmentation* operation) has been heuristically motivated with the emphasis on simplicity, data reduction, and comprehensibility of the resulting symbolic representations. Such transformations often result in a loss of information and a concomitant compromise in classifier performance.

On the other hand, structural approaches have significant advantages. Frequently the segmentation operation is, by the ingenuity of the designer, highly information preserving and simultaneously produces significant reduction in data dimensionality. Thus, while containing less than the full information available in the original sensor measurements, symbolic representations have proven to be sufficient for classification in a number of important areas [8,13].

A further benefit of the structural approach is the direct applicability of the theory of computing machinery [14] and relational structures [15] to the manipulation of the symbolic elements resulting from the segmentation operation. In addition, a symbolic modeling can easily allow for the introduction of other sources of knowledge into the decision process, particularly heuristically obtained knowledge.

Recall that a pattern representation presented to a structural classifier is in a symbolic form. When classification is defined by determination of an unknown objects "class" or "category", i.e. in a 1 of N sense, a structural classifier provides not only a determination of the class but also a *correspondence* or *match* between the constituent symbols of the unknown pattern representation and the constituent symbols of the chosen library element. The resulting match and class determination together provide an interpretation of the unknown object in the context of the library. Thus, structural techniques generalize the classification task and enhance the result.

Direct application of much of the theory of computing machinery and relational structures provides matches between symbolic representations which are constrained to be *crisp* [16,17]. That is, elements of two symbolic representations must be brought into exact correspondence with their relations completely preserved for a successful match to be declared. However, measurement errors in the sensor can cause changes in the observed symbolic representations which, in turn, degrade classifier performance.

Several authors have proposed techniques to allow for "inexact" matches [18,19,20]. These extensions include evaluation of some type of match quality function, usually in the form of a probability density function (or the related entropy function), thereby making use of the information contained in the symbolic representation. It has, however, been recognized that the loss of information in the segmentation operation can be a significant detriment to classifier performance [11,12]. Thus, a significant compromise in performance can result even when optimal use is made of the existing information.

1.2 Information Loss and Semantic Attributes

In order to minimize the loss of information, several authors have proposed the addition of semantic "attributes" to the symbolic description [8,21,22,23]. These attributes are designed to convey some portion of the parametric information contained in the original sensor data. This can result in a significant reduction in the loss of information during the conversion to a symbolic representation. Exploitation of the additional information is usually accomplished through an extension of symbolic methodologies via probabilistic techniques [21].

Much of the theory pertaining to the combined semantic–structural approach assumes the existence of a probabilistic model of the pattern generation process [22,24]. Often, assumptions regarding the statistical independence of the random variables making up the observation model are also implied [22,24,25,26,27]. In [23] the authors extend the independence assumption as to restrict the use of semantic information to the point that it is used only to resolve the decision of a structural classifier into *sub-classes*. Nonetheless, the results obtained using these approaches give an indication of the gains realizable through incorporating semantic (attribute) information [24,28].

This report presents a methodology in which symbolic representations are coupled with metric attributes in a way that makes weaker assumptions regarding statistical independence of the semantic and structural parameters. Such a quantitative approach to the structural problem is viewed as a

complete synthesis of statistical and structural methods of pattern recognition. A major contribution of this report is the demonstration of a pattern recognition system which fully exploits the relationship between structural and semantic information in the pattern representation.

1.3 Determination of Match Sense

In a structural pattern recognition system, the resulting match between elements of the unknown symbolic pattern representation and the symbolic pattern representation of a given library element can be viewed as a mapping from one set of symbols to the another. This mapping can be either injective or surjective or both. Furthermore, the image and pre-image of the mapping may cover arbitrary subsets of the symbolic representations of the unknown or library. The characteristics of the resulting mapping are collectively referred to as *match sense.*

The theory pertaining to matching of relational structures requires that a definition of match sense be provided prior to classification [19]. By considering all possible matches between the symbolic representations, the need for providing prior definition of match sense is eliminated. Under the classifier proposed here, prior restriction of match sense is possible but not necessary. Prior restriction of match sense may be done to reduce the amount of computation necessary to determine a match. Thus, by considering matching in this more general sense determination of an optimal match brings with it determination of the optimal match sense.

Determination of match sense is accomplished by adopting a view of the correspondence between two symbolic representations as being a partition of the union of the two structural descriptions. This view brings a natural sense of symmetry to the matching task which is not present under the mapping view of matching [29].

1.4 The Target Identification Application

The benefits of a structural approach to target identification, as well as the applicability of the described pattern recognition methodology, are demonstrated with an application to the target identification problem. Matching of two symbolic representations which represent radar targets is accomplished by assuming that the statistical properties of the measurement process carry over into the segmented representation of the radar object. By this it is meant that the distance function, which is minimized to produce a statistically optimal decision regarding target class, is also used for the match quality function to produce decisions regarding the elements of the segmented radar measurement series.

This distance function, as applied to elements of the symbolic representation, does not produce a statistically optimum decision except under a trivial special case. In this special case all information regarding the segmentation of the measurement series (all structural information) is discarded and the structural classifier is equivalent to the statistically optimal classifier. However, since the structural information is discarded at this point, no benefit can be derived from it.

Segmentation of the radar measurement vector is accomplished via a parametric estimation procedure. The chosen procedure is a modification of the Prony method for estimating the parameters of a function which can be modeled as a sum of exponential functions. Given the values of the function at discrete, equally spaced, intervals, the estimation procedure produces estimates of the exponential parameters. This estimation algorithm is applied to a series of frequency domain measurements of the "transfer function" of a radar object.

The estimation procedure exhibits a segmenting property in the range domain, thereby extracting the "scattering centers" of the radar object. A symbolic representation therefore consists of set of scatterers. The structural portion of such a symbolic representation is given by the existence of a given scatterer and the location of the given scatter, in range, with respect to the other scatterers. The semantic portion of the symbolic representation consists of the other parameters attached to each scatterer.

The Prony method, employed here, has been previously proposed for use in radar target identification [30]. In this research, the authors use a physically motivated exponential model for the time or range domain "ramp response". The Prony algorithm is therefore applied to a ramp response which is estimated using numerical methods applied to wire models of aircraft, or from empirically derived frequency domain radar object measurements. The objective in extracting the parameters of the exponential model is to estimate a set of "complex natural resonances" (CNR) corresponding to the given target. This research demonstrates how a given radar object can be accurately represented with a relatively small number

of complex numbers. Furthermore, in [30] the authors demonstrated the discriminating power of the resulting CNRs.

Application of the Prony method to the time domain, does not, however produce a segmentation of the radar measurement in the time domain. Therefore, there exists no geometric interpretation of the elements of the representation and the segmenting property of the estimation routine is unused. The discriminating power of the CNR representation is expressed in terms of correlation between the resulting time domain ramp responses of two different targets.

In the current application rejection of extraneous responses is accomplished by modifications to the match quality function in which the cost associated with rejection of scatterers is reduced. Thus the use of a structural method of pattern recognition allows for increased robustness, with respect to the presence of extraneous, unmodeled responses as is shown in Chapter 4.

The well-defined statistical modeling of the radar measurement process and the assumed extension of the statistical model to the symbolic representation suggests many of the extensions to the theory of relational graph matching made in this report. As mentioned previously, the match quality function is a result of the statistical modeling of the radar measurement process. The match quality function has a natural form which is "symmetric" (i.e. if A and B are two symbolic representations then the minimum distance correspondence from A to B is the same as from B to A). This, in turn, suggests the "partition" concept of matching and the automatic determination of match sense.

1.5 Summary

In order to motivate the methodology used here and fully describe the problem, the material in Chapter 2 presents the assumed radar system model and associated signal processing. The processing of the radar measurements is, in the current case, analogous to the segmentation operation. The parametric estimation algorithm is described and important properties are discussed.

In Chapter 3 a notational framework for a matching algorithm which possesses the required properties (determination of match sense and a relationship to statistical methods) is provided. In this chapter, notation for symbolic descriptions from previous authors [19] which is based on relational graph matching is adopted. In Chapter 3, qualifications placed on the relational portion of a structural description, which are considered necessary for complete fusion of structural and attribute information are discussed. The resulting properties of the relations are illustrated with an example using the notational framework.

In order to make full use of the semantic information contained in the symbol attributes, the existence of a meaningful sense of distance between elements of a symbolic representation is assumed. It is further assumed that this distance is defined, not only between the elements of two separate symbolic representations, but also between elements of a single symbolic representation. Chapter 3 continues with a precise definition of this distance, relating the distance between two sets of nodes to the distance between two symbolic representations.

Implications of such a model for the pattern generation process are made explicit by a comparison of this model with other models of structural description matching. Chapter 3 concludes with some notes concerning issues of searching for the optimal matches between structural descriptions and degenerate cases for distance functions.

In Chapter 4, application to the target identification problem is described in detail. For this problem, the radar measurement is segmented with the parametric estimation procedure described in Chapter 2. The interpretation of the parametric representation as a structural description required for classification is made explicit. This allows for the formal implementation of the algorithms described in Chapter 3.

The properties of the parametric representation of the radar measurement vector are used to define the required sense of distance between symbolic representations. The chapter concludes with results from experimental computer simulations which give an indication of the performance and benefits of a system which uses a structural method.

In the final chapter, the results of this research are summarized and the expected impact is outlined. Directions for future research are given.

Chapter 2

RADAR SYSTEM MODEL AND SIGNAL PROCESSING

In this chapter, a model for the radar measurement process is described. The model is based on a straight-forward approach to obtaining frequency diverse, continuous wave (CW), *coherent* [31] measurements of the backscatter from a radar target. The measurement process is described in some detail to support the notion that the required measurements can be made with an actual radar system.

For simulation and performance evaluation purposes, data obtained from a compact range is used [32]. The measurements made with a compact range are continuous wave, coherent measurements taken at a range of frequencies. By demonstrating the feasibility of obtaining coherent CW measurements by an actual radar system, this discussion justifies the use of measurements obtained from a compact radar range for simulation and performance evaluation.

In addition, a parametric estimation technique, which is to be applied to the measurements is described. The resulting parameter estimates are to be

used in place of the CW measurements for target identification. Thus, the resulting parameter estimates are referred to as a *parametric decomposition*. Using this technique, the continuous wave measurements are implicitly converted to an estimate of an "impulse response" or "range profile" of the illuminated target. Other salient properties of parametric decompositions are discussed.

2.1 Radar System Model

The envisioned radar platform is a stepped frequency radar system operating in approximately the HF band. This range of frequencies is chosen to correspond to the resonant region (wavelength approximately equal to object extent) for a given library of aircraft. Measurements taken from this library are used for simulation purposes in Chapter 4.

The radar is assumed to operate by transmitting a pulse which consists of a series of sub-pulses. Each sub-pulse is at a different frequency. A series of radar measurements are made in which each measurement corresponds to a different sub-pulse (Figure 2.1). The duration of each sub-pulse is determined such that elements of the radar measurement series $\{y_k\}_{k=0}^{N-1}$, represent "steady-state" values. In other words, the measurement series is equivalent to a sampling of the "transfer-function" corresponding to the given radar target (Figure 2.2). It is assumed that measurements taken with pulses of l cycles or more will accurately give such values.

It is also assumed that the number of sub-pulses is N, each of length $\frac{\tau}{N}$ seconds. In order to simplify the measurement process, all sub-pulses

are of equal length. Thus the minimum number of cycles per sub-pulse constraint applies only to the lowest frequency sub-pulse. The duration of the entire pulse must be sufficiently short such that coherency is preserved. For the experiments used this study, the number and duration of sub-pulses is chosen such that this is true. Coherency preservation is demonstrated by showing that the maximal error in the phase estimate of the final measurement with respect to the first measurement is negligibly small.

Individual sub-pulses are ordered in frequency such that the initial frequency is f_0 Hz and that subsequent sub-pulses are stepped in frequency by Δf Hz. Note that this ordering of the pulses is not required for the current application; sub-pulses can be sent in an arbitrary order and then re-ordered upon reception. The radar measurement corresponding to frequency f_k is denoted y_k for $0 \leq k \leq N-1$. The series of measurements can be represented in vector form as

$$\vec{\Upsilon} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-1} \end{bmatrix}. \tag{2.1}$$

Since the measurements are discrete in frequency, transformation of the series of measurements to the time (or range) domain via the Fourier transform implies periodicity in the range dimension. The unambiguous range, R, is determined by

$$R = \frac{c}{2\Delta f}. \tag{2.2}$$

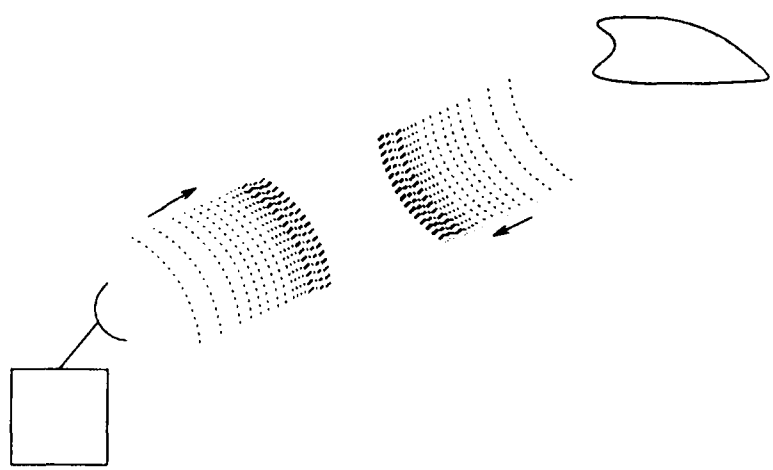


Figure 2.1: Stepped frequency radar system.

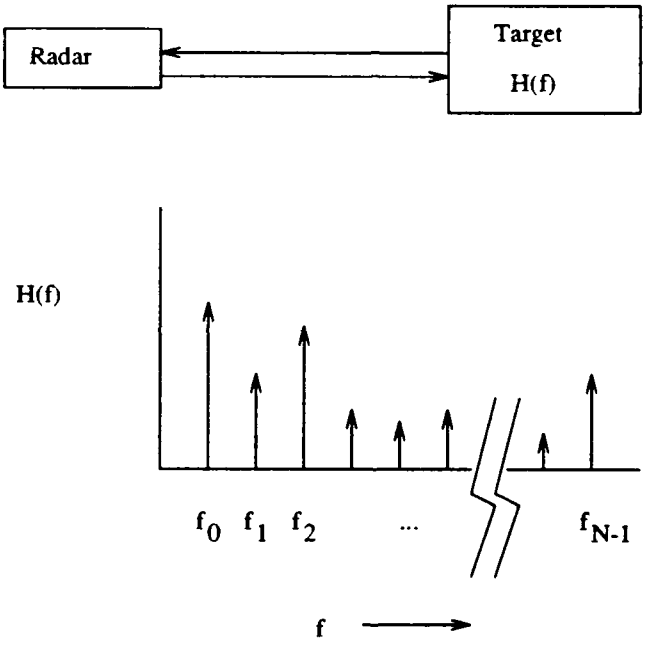


Figure 2.2: Stepped frequency radar system.

In a surveillance radar system, the search space is divided in the range dimension into a number of *range cells*. The range cell size, δr, is related to the time duration of the total pulse, τ, by

$$\delta r = \frac{c\tau}{2N}. \tag{2.3}$$

If the unambiguous range is chosen to be the range cell size then the initial frequency, f_0 and the frequency spacing are linked via the minimum number of cycles per sub-pulse, l as

$$f_0 = l\Delta f. \tag{2.4}$$

Using the parametric estimation techniques discussed below, the number of radar measurements, N, determines an upper bound on the number of *scattering centers* (Section 2.3.1) which can be extracted from the measurements. Thus, once appropriate numbers for N, l and δr are determined from the catalog of interest, the radar system parameters easily follow:

$$R = \delta r \tag{2.5}$$

$$\Delta f = \frac{c}{2\delta r} \tag{2.6}$$

$$f_0 = \frac{lc}{c\delta r} \tag{2.7}$$

$$\frac{\tau}{N} = \frac{2\delta r}{c}. \tag{2.8}$$

2.2 Prony-Based Parametric Estimation

A parametric estimation procedure is applied to the radar measurement series to provide reduction of data dimensionality and to segment the energy in a measurement vector into discrete elements which are approximately disjoint in the range domain. The parametric estimation procedure has other desirable properties as is discussed below and in [33]. The segmenting property of the estimation process is covered in Chapter 4.

The procedure used is an adaptation of the Prony method by R. Carriere. The Prony method is used for estimating parameters of exponential functions. Details of the procedure are given in [33]. In this section, the method is outlined and important properties of the technique are made explicit.

The technique used here is essentially a parametric spectral estimation technique [34]. In spectral estimation, a sample of a time signal is input to the algorithm and an estimate of the frequency spectrum results. Under our model, a series of frequency domain measurements of the "transfer function" of an object is input to the algorithm and an estimate of the resulting "impulse response" or "range profile" is produced.

2.2.1 The Frequency Domain Parametric Model

The model used for this estimation procedure is, in the frequency domain

$$y_k = \sum_{i=1}^{M} d_i p_i^k \qquad 0 \le k \le N-1. \tag{2.9}$$

where y_k is the k^{th} component of the measurement series (k is a subscript) and elements of set $\{p_i \mid 1 \leq i \leq M\}$ are referred to as the *poles* of the models and elements of the set $\{d_i \mid 1 \leq i \leq M\}$ are the corresponding *residues* (residue-pole pairs are referred to as *modes*). M is the *order* of the model. A parametric decomposition, Λ consists of the entire set of modes which, together, represent the measurement series, $\Lambda = \{(d_i, p_i) \mid 1 \leq i \leq M\}$. Examination of Equation (2.9) implies that the measurement series $\{y_k\}_{k=0}^{N-1}$ can be rewritten as a sum of a set of M series, $\left\{\{d_i p_i^k\}_{k=0}^{N-1} \mid 1 \leq i \leq M\right\}$. Thus, in the same way that the measurement series has a vector form, each mode of a parametric decomposition has a vector form. For example, a vector form, $\vec{v}$, of a mode (d, p), is given by

$$\vec{v} = d \cdot \begin{bmatrix} p^0 \\ p^1 \\ \vdots \\ p^{N-1} \end{bmatrix}. \tag{2.10}$$

Equation (2.9) then becomes

$$\vec{\Upsilon} = \sum_{i=1}^{M} \vec{v}_i. \tag{2.11}$$

2.2.2 The Range Profile

The modes of a given parametric decomposition have a range (time) domain interpretation in addition to the frequency domain interpretation given by Equation (2.9). The profile of a target corresponding to a given measurement series as a function of range, r, is

$$Y(r) = \sum_{i=1}^{M} \frac{d_i}{e^{-j\frac{2\pi r}{R}} - p_i} \quad -R/2 < r < R/2. \tag{2.12}$$

This equation can be derived by taking the inverse Fourier transform of Equation (2.9) under the assumption that the series for "stable" modes (modes with pole modulus less than 1) continue in the positive frequency direction indefinitely and that the series for "unstable" modes continue in the negative frequency direction indefinitely.

The energy from a given mode is centered in range about the range location of a mode, given by

$$x_i = -\frac{\arg(p_i)}{2\pi} R. \tag{2.13}$$

This range location is with respect to the "phase center" of the range cell. Since targets are not assumed to lie at any particular location within a range cell, the location of this reference point is arbitrary.

Operation of the estimation algorithm is now outlined, details are available in [33]. The algorithm begins by estimating a set of characteristic polynomial coefficients for the frequency domain series $\{y_k\}_{k=0}^{N-1}$ using the Yule-Walker Equations [34]. A "total-least-squares" algorithm with a singular value decomposition [35,36] is used to solve the Yule-Walker Equations. Only the M largest singular values of the augmented Hankel matrix are retained. This step is taken for "noise cleaning" purposes. The roots of the characteristic polynomial are derived via standard polynomial rooting techniques to produce a set of candidate p_i parameters. From the derived p_i parameters and the modeling Equation (2.9) the corresponding d_i are estimated using a least squares procedure. Since the Hankel matrix is of reduced rank, only the M largest energy modes are retained for classification. Furthermore, only the poles (the p_i parameters) are actually retained;

the corresponding d_i parameters are re-estimated following rejection of the low energy modes. In anticipation of the noise cleaning step, the initial estimate of model order is made somewhat higher than the expected number of scattering centers, M.

2.3 Properties of a Prony Parametric Decomposition

A parametric decomposition, as derived by the estimation technique described in Section 2.2, has a number of desirable features for target identification. By using a parametric decomposition instead of the full measurement series the sensed objects are represented by significantly fewer numbers of parameters. Furthermore, it can be shown that the variance of the location of a mode in the range domain is reduced (with respect to the variance of elements of the measurement series). Experimental data shows that there exists a strong relationship between elements of a parametric decomposition (through the impulse response) and object geometry [37]. The relationship to target geometry is demonstrated by comparison of down-range profiles to silhouettes of targets, and in an analytic way by the range domain *energy-localization* property of a parametric decomposition.

The parametric estimation procedure can yield significant reduction in data dimensionality. Recall the model order M determines the maximum number of modes which are extracted from the measurements. Then $2M$ (twice the model order) is the total number of complex parameters needed to represent the object with a parametric decomposition of M modes. This

number, $(2M)$ is usually chosen to be significantly less than N, the number of complex measurements.

If $2M$ is chosen to be less than N then it is not possible to find a set of d_i and p_i parameters such that Equation (2.9) holds for an arbitrary measurement series, $\{y_k\}_{k=0}^{N-1}$. In this case the d_i parameters in Equation (2.9) are determined by a least-squares estimate. The difference between the observed measured series, $\{y_k\}_{k=0}^{N-1}$, and the series $\{\sum_{i=1}^{M} d_i p_i^k\}_{k=0}^{N-1}$ is referred to as *modeling error*.

In addition to a reduction in data dimensionality, a further benefit of using this parametric estimation procedure is that the estimates of mode range location (the center of the energy) can be highly noise insensitive. In addition, separation of peaks in the range profile is not limited to a DFT (Discrete Fourier Transform) bin-width. This property is analogous to super-resolution of spectral components of a time signal using parametric spectral estimation techniques [33,34].

Due to the non-linear nature of the estimation process, the parametric decompositions also exhibit a number of undesirable characteristics under noise perturbations. In particular, a small perturbation in a measurement vector can result in an extreme change in the output parametric decomposition. Such wild variations in the parametric decomposition can frequently be attributed to the fact that the correspondence between noise-perturbed modes of a parametric decomposition and the "true" or "noiseless" modes is not easily determined.

In Appendix A a remedy to the correspondence problem based on the statistical characteristics of the measurement process is proposed. Using

this solution, an examination of the statistical properties of the parameters under noise perturbations can be made. The random nature of parametric decompositions is discussed in subsection 2.3.2.

2.3.1 Target Geometry and Energy-Localization

A primary motivation for using the parametric model described here is that the down-range profile, as specified by Equation (2.12), can exhibit a high correlation with target geometry. This feature is demonstrated by Figure 2.3[1]. This figure contains range profiles estimated from simulated radar measurements taken with a Concorde aircraft at two different aspect angles with respect to the radar.

The observed relationship between the geometric features of a target and the down-range profile can be summarized as follows. First, it is assumed that incident energy on the target is predominantly reflected from geometrically relevant portions of the target (edges, corners, engine inlets etc). The down-range profile represents energy from such "scattering centers", projected tangentially onto a line emanating radially from the radar to the target (the radar–target axis or line-of-sight vector).

Figure 2.3 also includes drawings of the target from which the radar measurements were made. These drawings are displayed at an angle with respect to the range axis of the range profile to indicate the angle which the target makes with the radar–target axis. Also included are vertical lines showing the assumed correspondence between elements of the target geometry and peaks of the down-range profile. Peaks of the down-range profile

[1]Thanks to R. Carriere and R. L. Moses

show a clear structural correspondence to the indicated scattering centers of the target. For the case given in this figure, the "projection" model works almost flawlessly. While such artwork cannot serve as an analytic tool for examining the relationship between object geometry and electromagnetic scattering, this discussion is offered as an intuitive justification for the use of parametric representations for target identification.

Examination of the range-domain model in Equation (2.12) shows that peaks in the down-range profile correspond to the modes of the frequency measurement series. Thus the range location of the peaks of the range domain response are easily derived from the parametric representation using Equation (2.13). The energy contained in a mode as well as the range extent of a mode are simple functions of the p_i and d_i parameters of the mode. The peaks in the down-range profile estimate are not necessarily required to follow an exact projection law with respect to target geometry in order to provide information sufficient for classification. All that is required is that different targets produce statistically different down-range profiles.

As is discussed in Chapter 4, the most important property of the parametric estimation procedure for application to structural pattern recognition is the way in which the energy in the measurement series is *segmented* in the range domain. Segmentation is seen here as being equivalent to *localizing* mode energy in the range dimension. In order to characterize the energy localizing property of the parametric estimation procedure, the relationship between pole modulus and the modes energy density is given.

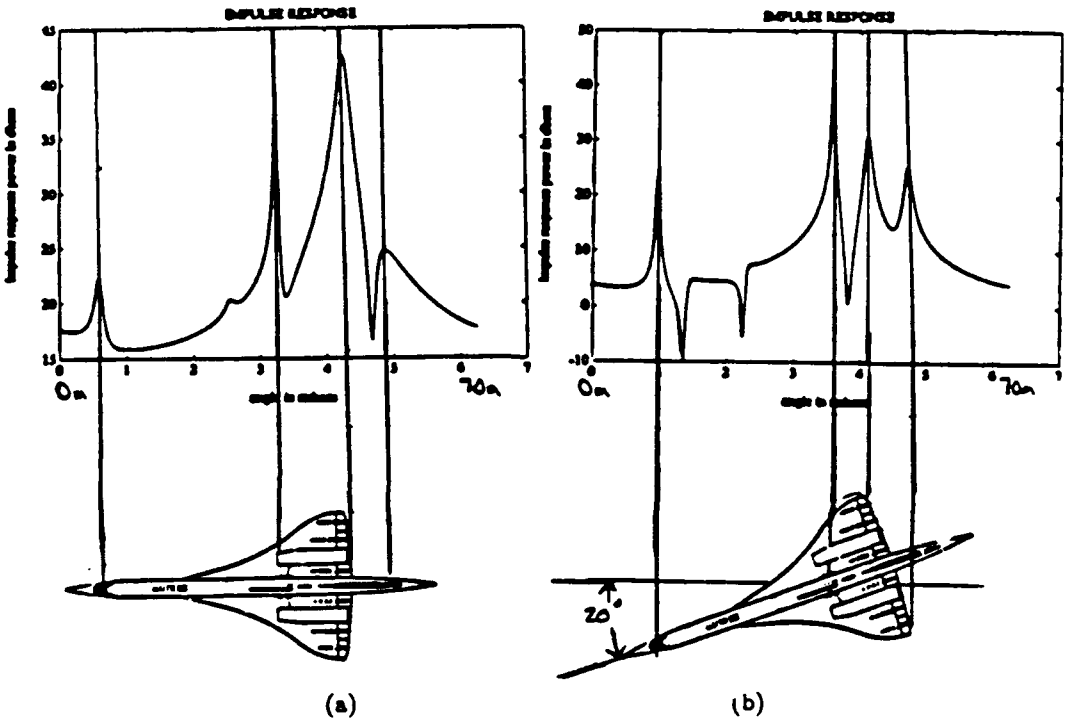


Figure 2.3: Concorde range response estimates with corresponding silhouettes made to scale. (a) At 0° Azimuth. (b) At 20° Azimuth.

In addition the relationship of pole modulus to the half-energy width of the mode in the range dimension is made explicit. These relationships are derived using algebraic manipulation and simple calculus.

For a pole, p of given modulus ρ and known range location, r, fix a sub-interval of the unambiguous range interval centered about r. Denote the proportion of the unambiguous range interval which is covered by this interval as β. Denote the proportion of the total energy contained in the pole over the sub-interval as E. Then by solving some simple integrals, a relation between the energy proportion, E, and the unambiguous range proportion, β is established:

$$\tan\left(\frac{\pi}{2}E\right) = \left|\frac{1+\rho}{1-\rho}\right| \tan\left(\frac{\pi}{2}\beta\right). \tag{2.14}$$

If the *energy density* of a mode is defined as $\frac{E}{\beta}$ then this function can be shown to be a non-increasing function of β. The energy density as a function of range proportion is shown for several values of pole modulus in Figure 2.4. Since it is non-increasing, this quantity achieves an upper-bound as $\beta \to 0$. This upper bound is determined by taking a limit:

$$\lim_{\beta \to 0} \frac{E}{\beta} = \left|\frac{1+\rho}{1-\rho}\right|. \tag{2.15}$$

In addition, if by defining the *half-power rangewidth*, γ, as the proportion of the unambiguous range interval spanned by the interval between points which are one half the peak energy of a given mode then:

$$\gamma = \frac{\sin^{-1}\left(\left|\frac{1-\rho}{\sqrt{\rho}}\right|\right)}{\frac{\pi}{2}} \quad \frac{2}{3+\sqrt{5}} \leq \rho \leq \frac{3+\sqrt{5}}{2}. \tag{2.16}$$

Values of ρ outside the given interval will produce half-power rangewidths which are greater than the unambiguous range interval. The pole modulus thus indicates the energy concentration of a mode in the range dimension. For example, a pole with modulus .95 will have a peak energy density of 39 and will achieve 90% of its energy in 10% of the unambiguous range interval. Furthermore, the half-power rangewidth is 3.2% of the unambiguous range interval. Peak energy density as a function of pole modulus is depicted in Figure 2.5.

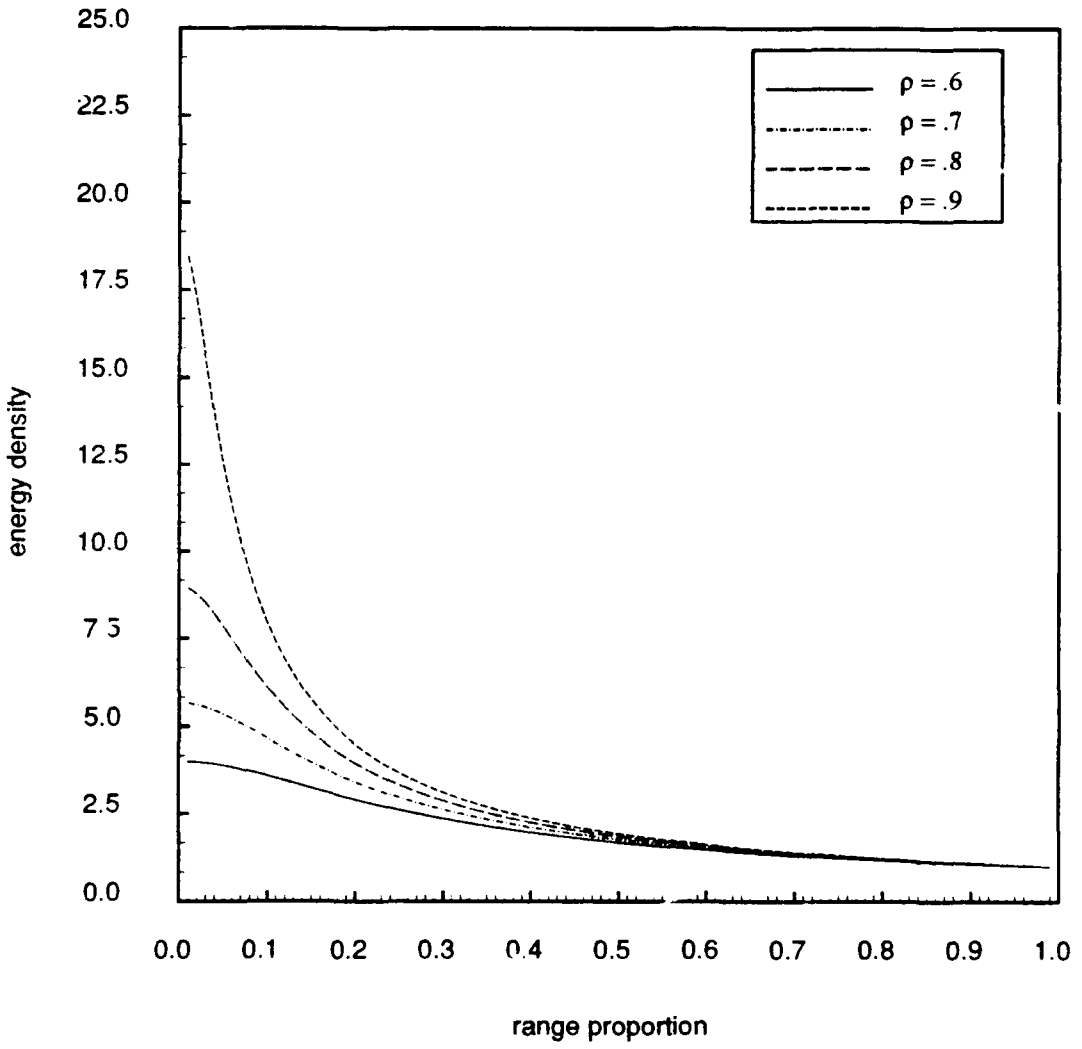


Figure 2.4: Energy density, $\frac{E}{\beta}$. vs range proportion, β, for modes with differing pole modulus.

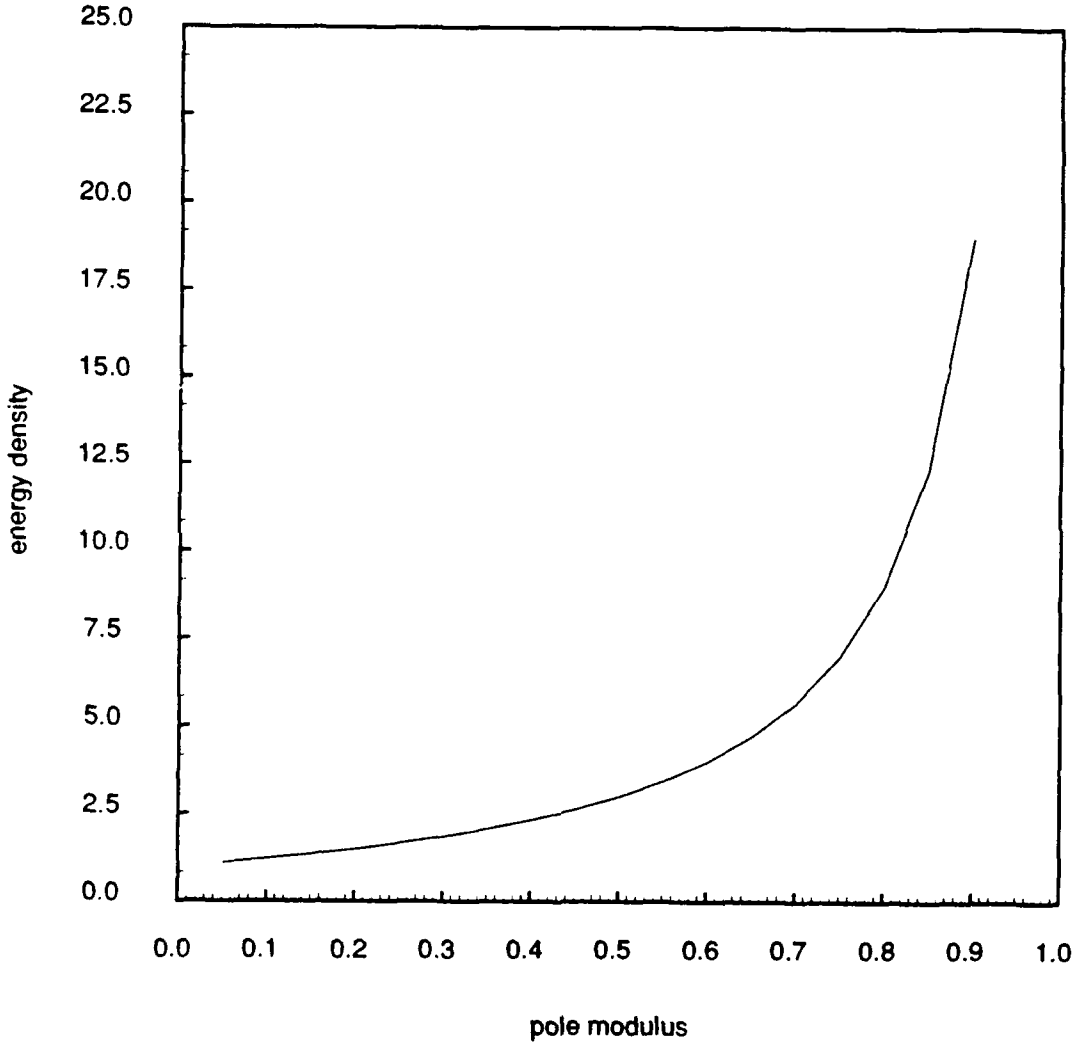


Figure 2.5: Peak energy density vs pole modulus.

2.3.2 Statistical Properties of the Parametric Decompositions

Noise perturbations induce variations in the parameters of the modes of a parametric decomposition. In addition the correspondence of noisy modes to their noiseless counterparts is blurred by noise perturbations. Variation in the pole location of a mode (the value of the p parameter) in the complex plane is illustrated in Figure 2.6. In this figure, the location of noisy poles in the complex plane is indicated by the numbers (1-5). Correspondence of the noise realization modes to the modes extracted from the noiseless measurements is accomplished by a method described in Section A.1 of Appendix A.

The noiseless pole locations, in the complex plane, and mode energy for the Concorde aircraft are given in Table 2.1. Comparison of the clusters of poles in Figure 2.6 with the parameters in Table 2.1 suggests that the variance associated with the location of a mode in the complex plane is mostly dependent upon the energy and pole modulus associated with the noiseless mode. Modes which contain high energy or which are highly localized appear to yield less variance in their location estimate. For example, estimates of the location of mode #1 have the "tightest" cluster due to the large energy of mode #1 coupled with the fact that the noiseless modulus for mode #1 indicates a peak energy density of 22 and is therefore fairly localized in range.

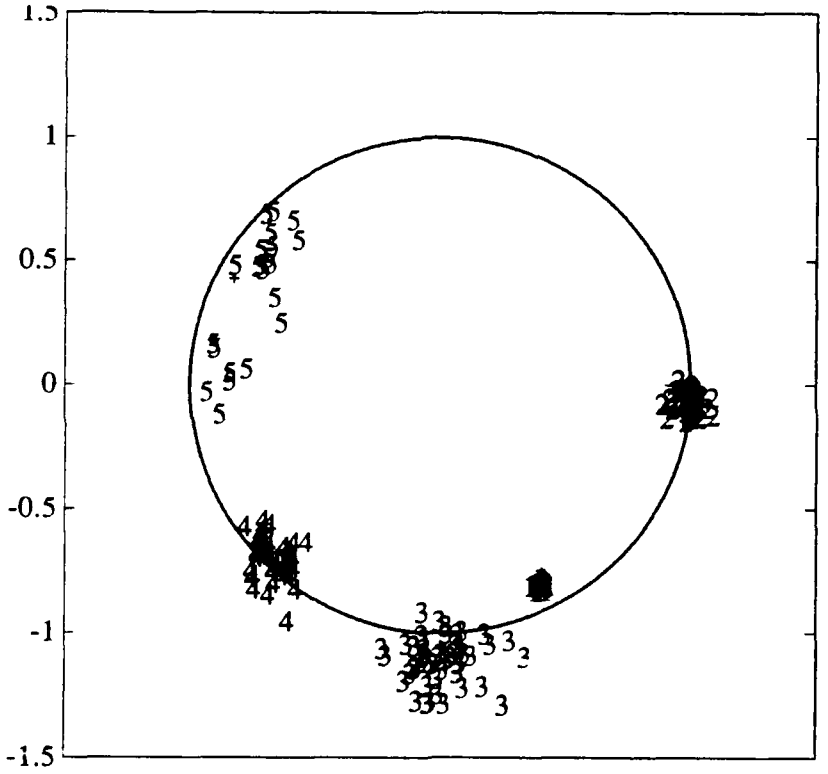


Figure 2.6: Concorde pole locations under noise perturbations.

Table 2.1: Noiseless mode parameters from simulated radar measurements for the Concorde aircraft.

Noiseless Mode	Mode energy	Pole angle	Pole modulus	Peak Energy Density
1	1932	-1.0911	0.9147	22.45
2	50	-0.0787	1.0240	84.33
3	36	-1.5674	1.0536	38.31
4	16	-2.3162	0.9320	28.41
5	6	2.6660	0.9278	26.70

Note that this is not always the case. Note that mode #3 has a higher noiseless energy and a higher peak energy density than mode #4. However, the pole locations indicated in Figure 2.6 for mode #3 are more dispersed than the pole locations for mode #4 thus contradicting the rule of thumb regarding parameter variance stated above. A likely reason for this paradox is that the indicated values for noiseless parameters are in error due to short data length constraints or unmodeled dynamics.

By using the categorization algorithm, the statistical properties of a given parametric decomposition can be determined. Under this method, a parametric decomposition of a noiseless measurement vector is formed. A number of parametric decompositions are formed of noise corrupted versions of the same vector. These parametric decompositions are then categorized with respect to their noiseless counterparts. The resulting parameters are then averaged to estimate mean and variance of the random parameters.

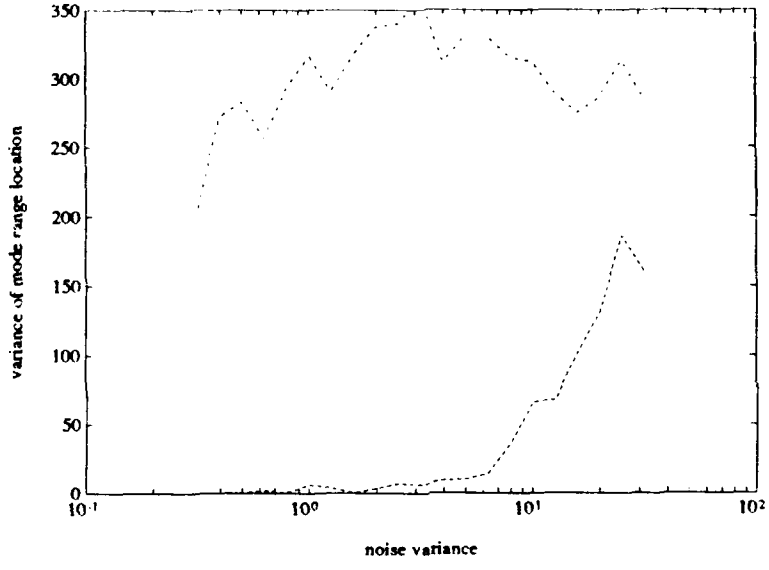
The variance associated with the mode range location is illustrated in Figure 2.7(a). In this figure, the variance of the range location for four different modes extracted from radar measurements of a Concorde aircraft is given as a function of the superimposed noise level. The noise environment

for this estimation procedure is that of zero mean white additive Gaussian noise. For each noise level, four hundred noise realizations are averaged to produce the estimates. Note the general trend of increasing variance with superimposed noise level. Variance estimates for the modes range location also appear to have an upper bound. This is due to the cyclic nature of the range location estimate. Also note that one of the modes occurs with essentially zero variance; the curve associated with this mode is obscured by the horizontal axis. This illustrates the variance reducing properties of the parametric estimation procedure for the mode range location parameter.

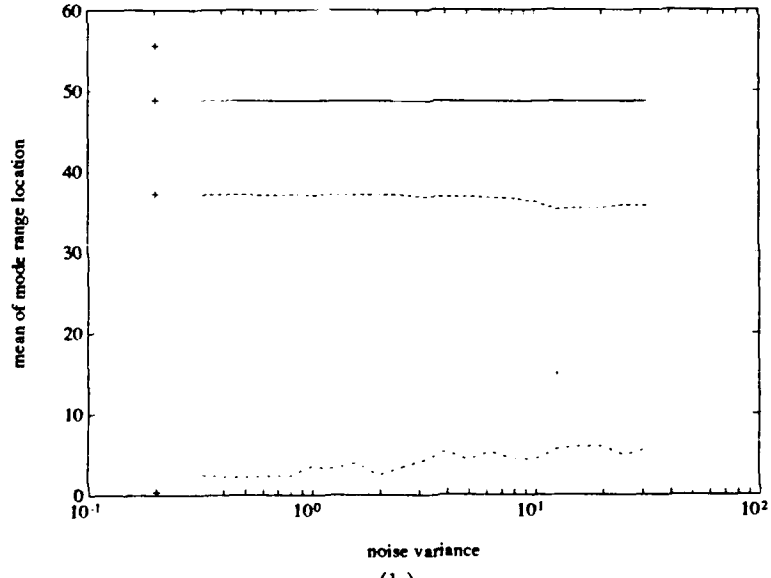
The curves in Figure 2.7(b) indicate estimates of the mean location in range of the four modes for the same case as described for Figure 2.7(a). The range location of the noiseless modes to which the noise realizations were mapped is indicated by the "+" marks in this figure. Note that the mode locations occur with little or no bias (the mean is approximately equal to the noiseless value). Indeed the non-zero estimates of mode bias appear to be due to an insufficient number of noise realizations for averaging.

Thus, the mode range location parameter (after categorization) can be characterized as being an unbiased random variable with variance which increases with the superimposed noise level. Similar comments hold for other parameters derived from a parametric decomposition.

The statistical properties of categorized modes suggests a method for estimating the likelihood function of a noisy parametric decomposition (conditioned on a given target class). Such an estimate can then be used to determine target class in a maximum likelihood sense. This method consists of two steps:



(a)



(b)

Figure 2.7: Estimated pole statistics versus noise variance. (a) Pole range location variance (m^2). (b) Pole range location mean (m).

1. For a given noisy parametric decomposition, categorize the modes of the decomposition into the modes of the known object representation.

2. Given the assumed noise level, determine the probability associated with the categorized parameters.

There are a number of difficulties associated with this method. To estimate the probability associated with perturbations of the parameters, the estimation of all moments of all parameters of the parametric decomposition is necessary. In addition, the algorithm used to determine the category of noisy modes determines a best *overall* match between two given parametric decompositions. Frequently, this involves the lack of a noiseless counterpart for certain modes of the noisy parametric decomposition (see Appendix A). For example, a match for mode #5 of Figure 2.6 occurred only 44% of the time. Therefore the estimation of the probability of a mode being present becomes necessary.

The difficulties associated with likelihood function estimation can be overcome by making a valid engineering approximation of the random parameters as being independent Gaussian random variables and by forming estimates of the probability of matching a given mode and incorporating these estimates into the likelihood estimate. Unfortunately, in addition to difficulties associated with the estimation of the conditional probability, there are a number of other reasons not to use this method.

The result of this estimation process is a measure of the probability that the given parametric decomposition is observed, under the assumed noise level, by radar measurements of the assumed object. However, the algo-

rithm which categorizes the noiseless modes produces a *distance* which indicates the degree of similarity that exists between the observed parametric decomposition and the assumed noiseless parametric decomposition. This distance can serve the same purpose as a likelihood function estimate.

The categorization algorithm, as implemented, is limited in the types of correspondences it is capable of making. It is possible that noise perturbations can cause one mode to be estimated in a noise realization where there are two in the noiseless estimate. The likelihood of such an occurrence depends upon the parameters of the noiseless modes. Likewise, it is possible that more than one mode can be estimated where there is only one in the noiseless parametric decomposition. These problems usually occur at higher noise levels. However, for complete generality, the probability of these events and the resulting parameter density must be estimated.

Categorization in this more general sense is capable of accounting for perturbations other than noise. Thus the use of symbolic matching algorithms is indicated for identifying targets which are represented by parametric decompositions. The fact that a meaningful sense of distance between arbitrary sets of poles can be defined (as will be discussed in Chapter 4) allows optimal matching, in a general sense, to be successfully implemented.

2.4 Radar System Parameters Example

As an example of how the key system parameters can be chosen for a given library of objects, consider a library of targets in which the maximum extent of the objects in the catalog is approximately 50 meters. The range cell size, δr should be chosen large enough such that the targets are likely to fall entirely within a range cell, and small enough such that individual scatterers can be separated. For our example the range cell size is chosen to be 75 meters so that

$$\delta r = 75 \Rightarrow R = 75 \Rightarrow \Delta f = \frac{c}{2\delta r} = 2\text{MHz}. \tag{2.17}$$

Thus, by choosing the minimum number of cycles per sub-pulse equal to 15 yields

$$l = 15 \Rightarrow f_0 = l\Delta f = 30\text{MHz}. \tag{2.18}$$

Suppose also that the number of peaks in the range profiles of the library targets is limited to 5. This determines the final model order, M, for the estimation procedure. As described in the subsection 2.2, a number of modes are rejected for noise-cleaning purposes. Supposing that 3 modes will be rejected following the initial estimation step, this yields an initial model order, before pole rejection, of 8. In order to estimate a model with 8 poles a minimum of 16 data points is necessary; $N = 16$. The overall pulse width is then

$$\tau = \frac{2N\delta r}{c} = 8\mu\text{sec}. \tag{2.19}$$

An object traveling at 225 meters/sec (about 500 mi/hr) will move 1.8 mm during the time between the commencement of the first sub-pulse and

the completion of the final sub-pulse. Thus, for this case, if the error in the estimated target velocity is 500 mi/hr then the error in the differential phase between the first and final measurements is .13°. In addition, the wavelengths corresponding to the transmitted frequencies vary from 10 to 5 meters. These wavelengths are within a factor of 10 of the extent of the targets in the library and, therefore, are within the resonant region for the given catalog. These radar system "dimensions" are assumed for the experiments described in Chapter 4.

Chapter 3

STRUCTURAL PATTERN DESCRIPTIONS AND ASSOCIATED DISTANCE MEASURES

In this chapter, a mathematical framework for use in matching the elements of two parametric decompositions is established. In this framework, the parametric decompositions of Chapter 2 are represented by *structural descriptions* [19]. The structural description formalism and the associated *relational graph* [8] formalism are useful for representing patterns symbolically and describing matching strategies. In Section 4.1 the relationship between parametric decompositions and the corresponding structural description is given. In this chapter the structural description notation is introduced. Furthermore, qualifying assumptions and a matching strategy useful for matching structural descriptions which represent parametric decompositions are presented.

The extension in [24] to the structural description formalism described in [19] is adopted here. The presentation of the matching formalism presented here is abstract in nature to stress the applicability of this theory to other areas. A more concrete presentation of the matching algorithm is given in Appendix A. Qualifying assumptions regarding the relational portion of the symbolic representation and the required distance measure illustrate the limitations of this theory. It is assumed that these qualifications are a necessary condition for the complete synthesis of symbolic and relational information and that they are natural in a variety of practical applications.

3.1 Structural Descriptions

A structural description, $\mathbf{D}$, is a 2-tuple, $\mathbf{D} = (\mathbf{P}, \mathbf{R})$ where $\mathbf{P}$ is a set of primitives, or nodes, of the structural description and $\mathbf{R}$ is a set of named N-ary relations over the set of nodes $\mathbf{P}$. Each element of $\mathbf{P}$ is a set of attribute/value pairs. In this way a node or primitive of a structural description can possess any number of properties. Thus, if $T = \{a_1, a_2, \ldots a_{N_A}\}$ is a set of node attributes and V_i, $1 \leq i \leq N_A$ is a set of allowed values for node attribute a_i then for any arbitrary $s \in \mathbf{P}$

$$s \subseteq \bigcup_{i=1}^{N_A} (a_i, v_i). \tag{3.1}$$

where $v_i \in V_i$.

Each element in the set $\mathbf{R}$ is represented by a 3-tuple,

$$r_k = (NR_k, R_k, f_{r_k}), \quad r_k \in \mathbf{R}. \tag{3.2}$$

where NR_k is the name of the relation and R_k is the set of N-tuples of elements of $\mathbf{P}$. This set lists groups of $\mathbf{P}$ which are related by the given relation. Finally, f_{r_k} is a function which assigns to each element of this set an attribute.

3.2 Relational Attributes

The basic assumptions which are employed concerning the relational portion of this framework are discussed below. These assumptions, listed as qualifications, form the characteristics of the class of structural relations considered here.

Qualification 1 *Relations are transitive. That is, there exists an operation, $\oplus$ which is applicable to relational atiributes such that:*

$$f_{r_l}[(s_i, s_k)] = f_{r_l}[(s_i, s_j)] \oplus f_{r_l}[(s_j, s_k)]. \tag{3.3}$$

Qualification 2 *All relations are binary and anti-symmetric.*

Hence,

$$(s_i, s_k) \in R_k \Rightarrow (s_k, s_i) \in R_k \tag{3.4}$$

and[1]

$$f_{r_l}[(s_i, s_k)] \oplus f_{r_l}[(s_k, s_i)] = 0. \tag{3.5}$$

[1]0, as used here, refers to "null" or "non-existent" and must be interpretable in the context of the attribute space. Relational attributes which are attached to involuted relations become part of the node attributes.

For example, attributes associated with a given relation could possibly represent a "directed distance" between nodes of a structural description in a particular dimension. Under such a scheme, the length from node s_i to node s_k is the "negative" of the length from node s_k to node s_i. Thus the term *anti*-symmetric.

A consequence of qualifications 1 and 2 is that there exists a variety of equivalent ways in which a given relation can be represented. Specifically, from a set of nodes connected in an arbitrary way by some relation in $\mathbf{R}$, a single node can be chosen to serve as the reference node for that specific relation. All nodes connected by such a relation need only be directly connected to the reference node. For example, in 2-dimensional Euclidean space (the relational attribute function evaluates to a 2-tuple), let $\mathbf{D} = (\mathbf{P}, \mathbf{R})$, with $\mathbf{P} = \{(\text{nodename}, 1), (\text{nodename}, 2), (\text{nodename}, 3), (\text{nodename}, 4)\}$, $\mathbf{R} = \{r_1\}$,
$r_1 = (\text{distance}, R_1, f_{r_1})$, and $R_1 = \{(1,2),(1,3),(3,4),(1,4)\}$ (for succinctness, nodes are referred to by their attribute values). Thus, we see from Figure 3.1(a) that $f_{r_1}[(1,2)] = (-1,2)$, $f_{r_1}[(1,3)] = (1,-3)$, $f_{r_1}[(3,4)] = (2,5)$, and $f_{r_1}[(1,4)] = (3,0)$.

Thus, by choosing node 2 as the reference node, the relation r_1 can be expressed in terms of pairings with node 2 as $R_1 = \{(2,1),(2,3),(2,4)\}$, (see Figure 3.1(b)) since

$$\begin{array}{lcl} f_{r_1}[(2,1)] + f_{r_1}[(1,2)] = (0,0) & \Rightarrow & f_{r_1}[(2,1)] = (1,-2) \\ f_{r_1}[(2,3)] = f_{r_1}[(2,1)] + f_{r_1}[(1,3)] & \Rightarrow & f_{r_1}[(2,3)] = (2,-5) \\ f_{r_1}[(2,4)] = f_{r_1}[(2,1)] + f_{r_1}[(1,4)] & \Rightarrow & f_{r_1}[(2,4)] = (4,-2). \end{array} \qquad (3.6)$$

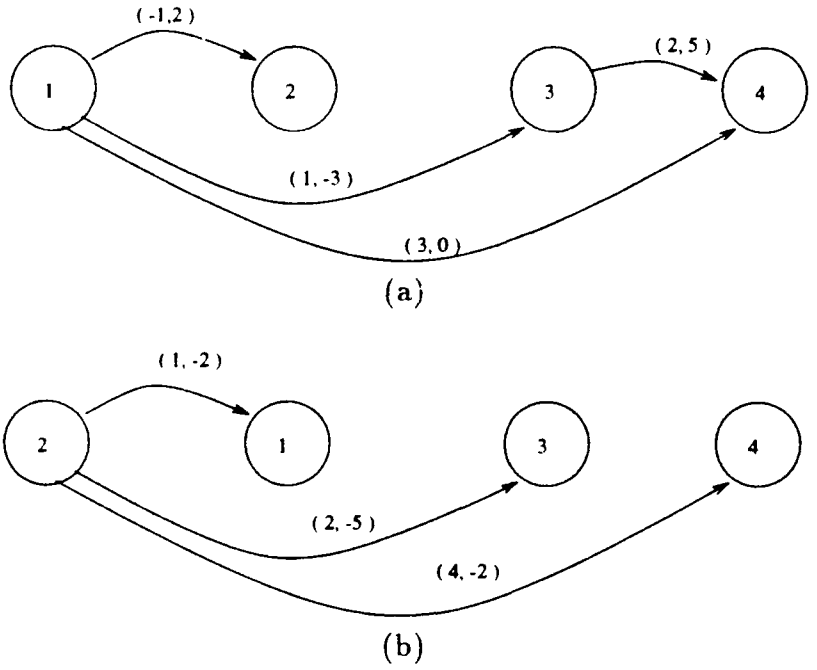


Figure 3.1: Structural Description with Transitive relations. (a) As originally specified. (b) Alternate specification with node 2 as a reference node.

Clearly any node can act as the reference node (including a fabricated reference node, i.e., a node which is appended to the structural description to act only as the reference node for one particular relation). It is assumed that such a reference exists for all relations in the given structural descriptions. In what follows, the reference node is denoted with an "R" in graphical depictions of structural descriptions.

The qualifications given above have consequences for the matching of structural descriptions. In particular, for a given match between the nodes of two structural descriptions the distances may be evaluated as a function of the registration (i.e., offsets, scalings, quadratic expansions etc.) between the relational attributes of two structural descriptions. As a result, the "optimum" set of registration parameters, i.e., those which best resolve the discrepancies between the relational attributes of the two structural descriptions, may be calculated using the distance function.

As an example, consider the trivial case in which each of the two structural descriptions, $\mathbf{D_A}$ and $\mathbf{D_B}$ consist of a single node. The only possible non-trivial correspondence between $\mathbf{D_A}$ and $\mathbf{D_B}$ is that which matches the single node in the observed description to the single node in the candidate description. In this case, relations (and therefore relational attributes) only exist from the singular nodes of the structural description to the fabricated reference nodes. Discrepancies in relational attributes can be resolved with only a single relational offset parameter in each dimension.

Consider now a second example in which $\mathbf{D_A}$ contains three nodes, as in Figure 3.2(a), and $\mathbf{D_B}$ has two nodes as in Figure 3.2(b). Under the correspondence between the two structural descriptions consisting of

$\{A_1 \leftrightarrow B_1, A_2 \leftrightarrow B_2, A_3 \leftrightarrow \emptyset\}$ all relational attributes can be brought into exact agreement with a relational offset of -1 unit in the "x" direction and +2 units in the "y" direction for the candidate structural description. Under a different correspondence, $\{A_1 \leftrightarrow B_1, A_3 \leftrightarrow B_2, A_2 \leftrightarrow \emptyset\}$, both an offset (-1.9,-1) and a subsequent scaling (multiplying the relations by (10,4)) of the relations of the candidate structural description are necessary to bring the relational attributes into exact correspondence.

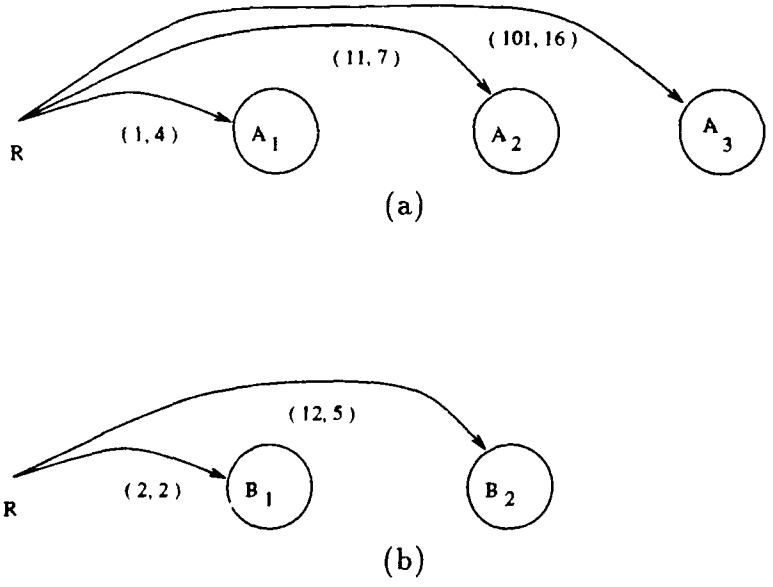


Figure 3.2: Matching of structural descriptions with transitive relational attributes. (a) Structural description $\mathbf{D_A}$. (b) Structural description $\mathbf{D_B}$.

Clearly, as the correspondence between the nodes of the structural descriptions becomes more complex, greater numbers of registration parameters must be added to bring relational attributes into exact equivalence.

By including sufficient numbers of registration parameters, full variability of relational attributes is allowed.

For many applications, problem domain knowledge may be used to limit the range of registration parameters. This is done by either explicitly stating a range of values for valid registration parameters or by simply excluding certain parameters and allowing no variability. In such cases, relational attributes cannot necessarily be brought into exact equivalence and a "best fit" solution must be adopted. This is accomplished by computing registration parameters which minimize an "inter-structural-description" distance function of the type discussed below. The formulation of the distance function between structural descriptions given here stresses the interdependencies between the relational and the symbolic (node attributes) components of the structural description.

3.3 Nodal Attributes and Inter-Structural-Description Distance

For certain applications, especially those in which statistical variations may not be significant, node attributes may be developed without regard for description comparison. Indeed, when node attributes have only an absolute interpretation, matching can only be precisely defined in a "crisp" sense. However, the approach followed here is one in which the "quality" of the match may be evaluated and used to advantage. Clearly, this measure of quality must be defined in terms of the relations between matched nodes. Hence, differences in attributes attached to relations between nodes influence the quality of a candidate match through registration parameters. The necessity of measuring the quality of a match leads to the final qualification on the structural descriptions:

Qualification 3 *For any two structural descriptions, there exists a meaningful measure of distance between subsets of nodes extracted from these descriptions. In addition, this distance is a function of the registration parameters.*

To make the notion of the distance measure precise, let two structural descriptions $\mathbf{D_A} = (\mathbf{P_A}, \mathbf{R_A})$ and $\mathbf{D_B} = (\mathbf{P_B}, \mathbf{R_B})$ be given by:

$$\begin{array}{ll} \mathbf{P_A} = \{s_{A_1}, s_{A_2}, \ldots, s_{A_{NA}}\} & \mathbf{P_B} = \{s_{B_1}, s_{B_2}, \ldots, s_{B_{NB}}\} \\ \mathbf{R_A} = \{r_{A_1}, r_{A_2}, \ldots, r_{A_{Nr}}\} & \mathbf{R_B} = \{r_{B_1}, r_{B_2}, \ldots, r_{B_{Nr}}\} \\ r_{A_i} = \left(NR_i, R_{A_i}, f_{r_{A_i}}\right) & r_{B_i} = \left(NR_i, R_{B_i}, f_{r_{B_i}}\right) \end{array} \tag{3.7}$$

where NA and Nr denote the number of nodes and relations, respectively, making up description $\mathbf{D_A}$.

The distance between a set of nodes from structural description $\mathbf{D_A}$ ($A \subseteq \mathbf{P_A}$) and a set of nodes from structural description $\mathbf{D_B}$ ($B \subseteq \mathbf{P_B}$), is denoted as $d_s(A,B)(rp_1, rp_2, \ldots, rp_{\mathrm{nrp}})$, where the registration parameters between node set A and node set B are denoted as rp_i, or simply as $d_s(A,B)(\vec{r})$ (where $\vec{r}$ is understood to represent a registration parameter vector, $[rp_1, rp_2, \ldots, rp_{\mathrm{Nrp}}]^T$). The distance $d_s(A,B)(\vec{r})$ must exhibit the properties:

$$d_s(A,B)(\vec{r}) \geq 0. \tag{3.8}$$

$$d_s(A,B)(-\vec{r}) = d_s(B,A)(\vec{r}). \tag{3.9}$$

Furthermore it is required that for any set of registration parameters, $\vec{r}$

$$d_s(A,\emptyset)(\vec{r}) = d_s(\emptyset,A)(\vec{r}). \tag{3.10}$$

Thus the *inter-node-set* distance, $d_s(A,B)$, must include the *nil-map* cost $d_s(A,\emptyset)(\vec{r})$ as a special case. A nil-map may occur when a particular node or set of nodes does not have a corresponding match in the other structural description.

The above definition naturally leads to a formulation of distance between two structural descriptions. In particular, suppose there exists a partition of the nodes of two structural descriptions $\mathbf{D_A}$ and $\mathbf{D_B}$, i.e., $\mathcal{A} = \{A_i \mid 1 \leq i \leq \mathrm{NA}\}$, $\cap\mathcal{A} = \emptyset$, $\cup\mathcal{A} = \mathbf{P_A}$ and $\mathcal{B} = \{B_i \mid 1 \leq i \leq \mathrm{NB}\}$, $\cap\mathcal{B} = \emptyset$, $\cup\mathcal{B} = \mathbf{P_B}$ (partitions of $\mathbf{P_A}$ and $\mathbf{P_B}$ are required to be mutually exclusive and collectively exhaustive). Furthermore, $\mathcal{M} : \mathcal{A} \mapsto \mathcal{B}$ is a binary one-to-one mapping from the elements of $\mathcal{A}$ to the elements of $\mathcal{B}$ (all unmapped elements of the partitions are implicitly "nil-mapped"). Then the

distance between the two partitions $\mathcal{A}$ and $\mathcal{B}$ is:

$$d_p(\mathcal{A},\mathcal{B}) = \min_{\mathcal{M}:\mathcal{A}\rightarrow\mathcal{B}} \left\{ \min_{\vec{r}} \left\{ \sum_{A_i \in \mathcal{M}^{-1}(\mathcal{B})} d_s(A_i, \mathcal{M}(A_i))(\vec{r}) + \sum_{B_i \notin \mathcal{M}(\mathcal{A})} d_s(\emptyset, B_i) + \sum_{A_i \notin \mathcal{M}^{-1}(\mathcal{B})} d_s(A_i, \emptyset) \right\} \right\}. \quad (3.11)$$

The distance between two structural descriptions, $\mathbf{D_A}$ and $\mathbf{D_B}$ is then defined as the minimum over all partitions of $\mathbf{P_A}$ and $\mathbf{P_B}$ as :

$$d(\mathbf{D_A}, \mathbf{D_B}) = \min_{\{\mathbf{Part}(\mathbf{P_A})\,\mathbf{Part}(\mathbf{P_B})\}} \{d_p(\mathcal{A},\mathcal{B})\} \quad (3.12)$$

where the set of all partitions of an arbitrary set $\mathbf{U}$ is denoted as $\mathbf{Part(U)}$. The order in which this minimization is accomplished is not important.

The partitions and the inter-partition mapping resulting from the above minimization do not necessarily constitute a relation between the nodes of the two structural descriptions which is constrained to be either one-to-one or onto. Indeed, the resulting correspondences may be homomorphic, isomorphic, monomorphic, sub-graph isomorphic, etc.

The lack of a pre-defined matching sense allows the matcher to make correspondences from groups of nodes to groups of nodes, thereby providing the potential to account for differing segmentations of the same pattern. Minimizing the inter-structural-description distance implicitly determines the sense of the correspondence between the nodes of the two structural descriptions. Since the sense of the optimum mapping is determined by the matcher itself, the only reason to restrict the match sense *a priori* is to reduce computational complexity.

3.4 Comparison with Other Models

In other models of structural matching, observed nodal and relational attributes are considered to be realizations of random variables which are modeled as either mutually independent or (in the case of relational attributes) are modeled as being dependent only upon the existence of the underlying graphs end nodes [22,27]. Such models are referred to as Independent Random Attribute (IRA) models.

By their definition, IRA models do not consider contextual information. In other words, the contribution to the match quality function from a perturbation to a particular nodal or relational attribute does not depend upon any other perturbations to any of the other nodal or relational attributes. For cases in which this independence assumption is a valid approximation, considering dependencies among the constituent random variables results in a more accurate assessment of match quality. Furthermore, the use of registration parameters to determine match quality implicitly introduces a dependence relation among the random variables.

As an illustration to this property, consider the two structural descriptions depicted in Figure 3.3. In this example, an observed structural description (Figure 3.3(a)) is assumed to be congruent to the prototype (Figure 3.3(b)) with the exception that the relational attribute between the two nodes has been modified by a vector, $\Delta\mathbf{r}$, of attribute perturbations. The assumed correspondence between the nodes from the two structural descriptions is $A_1 \leftrightarrow B_1, A_2 \leftrightarrow B_2$.

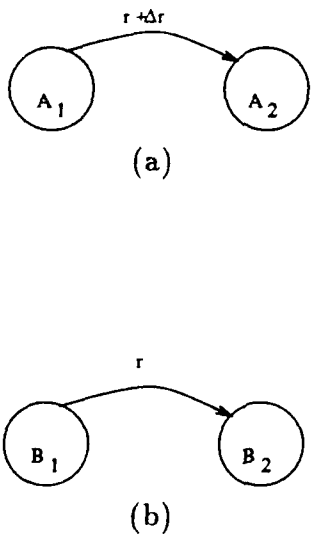


Figure 3.3: Contextual aspects of matching with registration parameters versus IRA models. (a) An observed structural description. (b) A prototype structural description.

In an IRA model the change in the match quality function, due to the perturbation of the relational attribute, is a function of only the observed and prototypical relational attributes,

$$\Delta d_{\mathrm{IRA}}(\mathbf{D_A}, \mathbf{D_B}) = f(\mathbf{r}, \mathbf{r} + \mathbf{\Delta r}) \tag{3.13}$$

Under a registration parameter model, a perturbation to any of the attributes of the noiseless structural description causes, in general, a change to the optimized values of the registration parameters. Therefore, the change in the inter-structural-description distance is a function of all nodal and relational attributes from both structural descriptions as

$$\Delta d(\mathbf{D_A}, \mathbf{D_B}) = f(\mathbf{r}, \mathbf{r} + \mathbf{\Delta r}, \mathbf{P_A}, \mathbf{P_B}) \tag{3.14}$$

Hence, the use of registration parameters represents an inclusion of dependence relations among the constituent random variables of an underlying random graph model. Such a model is applicable for matching measured realizations to library prototypes where variations of the measured realizations can be accurately described by variations in registration parameters.

The task of searching for a minimum map or match correspondence can be extremely time consuming, even for structural descriptions with very few nodes. In the following sections ways in which searching can proceed in an efficient way are described. These methods use the properties of the relational portion of the structural descriptions to reduce the number of partitions and relations which are considered.

3.5 Searching for the Optimal Match Between Structural Descriptions

Searching for the minimum distance partition and inter-partition mapping as suggested by Equation (3.11) and (3.12) appears to require a four step iterative process:

1. Form candidate partition of $\mathbf{P_A}$, ($\mathcal{A}$).
2. Form candidate partition of $\mathbf{P_B}$, ($\mathcal{B}$).
3. Form candidate mapping $\mathcal{M}$ between elements of $\mathcal{A}$ and elements of $\mathcal{B}$.
4. Compute the resulting distance function, minimize with respect to relational offset parameters.

In what follows, it is shown how the first three steps can be combined into a single step.

Consider the result of a formation of arbitrary partitions of the $\mathbf{P_A}$ node set along with arbitrary partitions of the $\mathbf{P_B}$ node set and a one-to-one mapping which links elements of the two partitions. Such a correspondence between the two node sets has the property that it can map elements of the node set to nil, or groups of elements of the node sets to nil, or elements of one node set to elements or groups of elements of the other node set and vice-versa.

Next, consider an arbitrary partition $\mathcal{P}$ of the combined sets $\mathbf{P_A} \cup \mathbf{P_B}$. Elements of $\mathcal{P}$ may contain a single element of $\mathbf{P_A}$ or a single element of $\mathbf{P_B}$

or subset of only $\mathbf{P_A}$ or $\mathbf{P_B}$ or they may contain elements of both $\mathbf{P_A}$ and $\mathbf{P_B}$. Elements of $\mathcal{P}$ are interpreted in the following way. Those elements of $\mathcal{P}$ which contain only elements of $\mathbf{P_A}$ or only elements of $\mathbf{P_B}$ are interpreted as mapping those elements (or sets of elements) to "nil". Those elements of $\mathcal{P}$ which contain elements of both $\mathbf{P_A}$ and $\mathbf{P_B}$ are interpreted as being a mapping from the portion of the partition element from $\mathbf{P_A}$ to the portion of the partition element from $\mathbf{P_B}$. Thus an equivalence is formed between partitions of $\mathbf{P_A} \cup \mathbf{P_B}$ and the formulation of a correspondence. Steps 1 through 3 of the search process can now be collapsed into a single step: Form a candidate partition $\mathcal{P}$ of $\mathbf{P_A} \cup \mathbf{P_B}$.

The distance function between two structural descriptions, discussed in the previous section, is now described using a simplified notation. Recall that an arbitrary partition of $\mathbf{P_A} \cup \mathbf{P_B}$ is denoted by $\mathcal{P}$. Let π denote an arbitrary element of $\mathcal{P}$, π_A denote the component of π in $\mathbf{P_A}$ and π_B denote the component of π in $\mathbf{P_B}$ so that $\pi_A = \pi \cap \mathbf{P_A}$ and $\pi_B = \pi \cap \mathbf{P_B}$. Then the distance between structural descriptions $\mathbf{D_A}$ and $\mathbf{D_B}$ is

$$d(\mathbf{D_A}, \mathbf{D_B}) = \min_{\mathcal{P}=\mathbf{Part}(\mathbf{P_A} \cup \mathbf{P_B})} \left\{ \min_{\vec{r}} \left\{ \sum_{\pi \in \mathcal{P}} d_s(\pi_A, \pi_B)(\vec{r}) \right\} \right\} \tag{3.15}$$

Enumeration of candidate partitions (or correspondences) can be accomplished via a recursive procedure. Details of the algorithm employed for simulation and performance evaluation are available in Appendix A.

3.6 Incorporation of Relational Constraints

Significant reductions of the number of candidate partitions may be made by considering only those partitions which are relationally "consistent." Correspondences between the nodes of two structural descriptions may be generated from other correspondences by recursively pairing elements of a given correspondence. Using this technique, the given input partition has associated with it a vector of optimizing values for the registration parameters at each level of recursion. Similarly, each potential pairing of elements of the input partition has associated with it a vector of optimizing values for the registration parameters.

The optimizing values of the registration parameters for a potential pairing of elements of the partition are defined by the corresponding inter-node-set distance function. For a given pair of elements from the input partition $\pi^1, \pi^2 \in \mathcal{P}$ this distance function is $d_s(\pi^1_A \cup \pi^2_A, \pi^1_B \cup \pi^2_B)(\vec{r})$. Optimizing values for the registration parameter vectors for the input partition and each potential pairing are assumed to be input with the partition.

By comparing the optimizing values of registration parameters of a potential pairing with those of the input partition, a decision can be made regarding the "relational consistency" of the potential pairing. This concept is illustrated by an example. Consider the structural descriptions depicted in Figure 3.4. An "offset" registration parameter is defined as the amount by which reference nodes of two structural descriptions are translated with respect to each other. In this example discrepancies in relational parameters are resolved with only an "offset" registration parameter. Assume that

the optimizing value of this registration parameter associated with an input correspondence (partition) of $\{(A_1, B_1), (A_2), (A_3), (A_4), (E_2), (B_3), (B_4)\}$ is 0. Optimizing values for some of the potential pairings of elements from this partition are:

Potential pairing	Optimizing offset
(A_2, B_2)	0
(A_3, B_3)	0.25

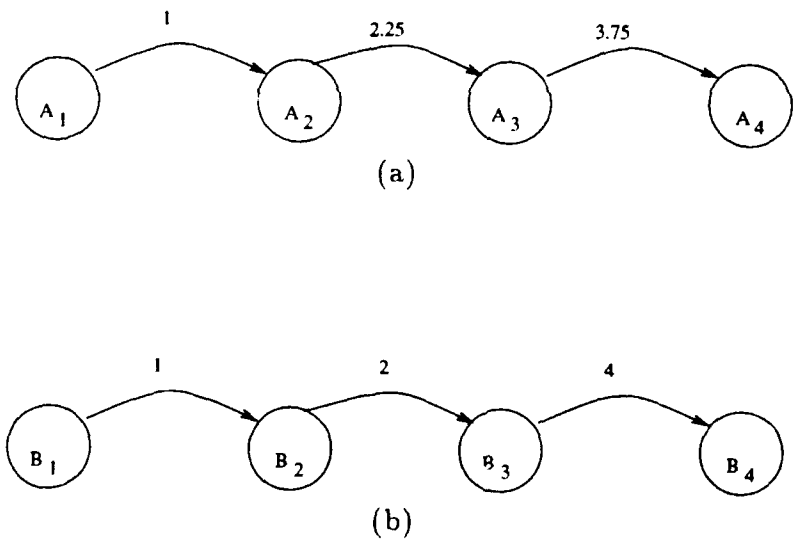


Figure 3.4: Relational consistency in matching. (a) An observed structural description with one node out of place. (b) A prototype structural description.

Clearly, the pairing (A_2, B_2) is relationally consistent with the input correspondence. The pairing (A_3, B_3) may or may not be relationally consistent depending upon how the registration parameters are compared. In the following section, an example of how relational consistency can be defined by changes to the match distance function.

3.7 Metric Inter-Node-Set Distances

Recall that inter-structural-description distance is simply the inter-partition distance minimized over all possible partitions of the joint node set of the two structural descriptions. The inter-partition distance is a sum of inter-node-set distances. Thus if the node set distance is a metric, i.e., if

$$d_s(A_i \cup A_k, B_i \cup B_k)(\vec{r}_c^*) \leq d_s(A_i, B_i)(\vec{r}_{si}^*) + d_s(A_k, B_k)(\vec{r}_{sk}^*) \qquad (3.16)$$

with $A_i, A_k \subseteq \mathbf{P_A}$, $B_i, B_k \subseteq \mathbf{P_B}$, and $A_i \cap A_k = \emptyset = B_i \cap B_k$, and where where $\vec{r}_c^*$, $\vec{r}_{si}^*$ and $\vec{r}_{sk}^*$ are optimizing values for the registration parameter vectors then the distance associated with any non-trivial partition of the node sets $\mathbf{P_A}$ and $\mathbf{P_B}$ is not increased by joining elements (forming the union of their node sets) of the partition.

A global minimum is therefore obtained by the partition consisting of only one element $\mathbf{P_A} \cup \mathbf{P_B}$. That is, the optimal match maps the union of all elements of $\mathbf{P_A}$ to the union of all elements of $\mathbf{P_B}$.

Therefore distance functions which satisfy the metric property allow (or force) the consideration of "entire-structure" matches as would be produced by a purely statistical pattern recognition algorithm. On the other hand, metric distance functions also preclude the possibility of extracting information regarding the correspondence between nodes of two structural descriptions. This characteristic is especially unfortunate since extracting the information contained in a correspondence between two structural descriptions is one of the primary objectives of structural pattern recognition.

It is interesting to note that this relationship between distance functions satisfying the triangle inequality and matching in a non-binary sense has

been previously referred to in [25]. There, the authors conclude that the distance function should not, in general, be a metric. Conversely, it is easy to see that if the inter-node-set distance is "too far" from being a metric function, there is no advantage of considering any correspondences other than one-to-one matches between nodes. In [25], the authors discuss a distance function for matching structural descriptions which allows for the more generalized type of matching considered here.

Two modifications to the inter-structural-description distance are proposed to induce non-trivial matches between the structural descriptions. The first of these serves to reduce the distance associated with matches having large numbers of correspondences between nodes. The second reduces the distance associated with matches having large numbers of nil-mapped nodes. Each of these modifications consists of a slight modification to the inter-structural-description distance function. The extent of these modifications may be viewed as an expression of prior knowledge concerning the type of matches which are to be made.

The modification designed to emphasize matches between nodes of two structural descriptions is referred to as *link normalization*. A link is defined as an element of the correspondence in which at least one node of the observed structural description is mapped to at least one node from the prototype structural description. Elements of the correspondence which map a node or a group of nodes to nil are not links. Likewise, elements of the correspondence which map more than one node of the observed structural description to one or more than one node of the prototype structural

description are counted as a single link regardless of the number of nodes involved.

In this modification, the distance associated with a given correspondence (or partition of $\mathbf{P_A} \cup \mathbf{P_B}$) is normalized by a factor dependent upon the number of links which are implied by the given correspondence. The number of links for a particular correspondence, $\mathcal{P}$ is denoted $\mathbf{L}(\mathcal{P})$. The modified distance function is the previously defined distance, divided by 1 plus the link-normalization-factor, α, times the number of links.

The second modification of the distance function considered involves the weighting of terms corresponding to nil-maps. This modification is intended to decrease the value of the distance function associated with correspondences involving nil-maps and, hence, cause such correspondences to be chosen with greater frequency. In this way problems associated with missing or extraneous in the prototype structural description are lessened. This factor is termed the *nil-map propensity*, ϕ.

By incorporating these two factors, the general form for the distance between two structural descriptions may be written as

$$d(\mathbf{D_A}, \mathbf{D_B}) = \min_{\mathcal{P}} \left\{ \min_{\vec{r}} \left\{ \frac{1}{1 + \alpha \mathbf{L}(\mathcal{P})} \times \left[\sum_{\substack{\pi \in \mathcal{P} \\ \pi_A \neq \emptyset, \pi_B \neq \emptyset}} d_s(\pi_A, \pi_B)(\vec{r}) + \phi \left(\sum_{\substack{\pi \in \mathcal{P} \\ \pi_A = \emptyset}} d_s(\emptyset, \pi_B) + \sum_{\substack{\pi \in \mathcal{P} \\ \pi_B = \emptyset}} d_s(\pi_A, \emptyset) \right) \right] \right\} \right\}. \quad (3.17)$$

Notice, from Equation (3.17), that an increase in the link-normalization-factor, α, will result in matches with greater correspondence between nodes of the two structural descriptions. An increase in the nil-map-propensity,

ϕ will result in matches with greater numbers of nil maps. Thus, in the sense that increasing the link-normalization factor forces greater complexity in the resulting matches, the resulting inter-node-set distance function is claimed to be "less metric".

Chapter 4

APPLICATION TO RADAR OBJECT IDENTIFICATION

An example of how the framework for structural descriptions can be applied to a problem involving the symbolic processing of waveform sources is now presented. The problem is that of identifying an airborne object from its real aperture radar responses. The object is assumed to be detected and a measurement vector (as discussed in Chapter 2) is assumed to be available.

Such problems have been approached from a statistical point of view previously [32]. The desired tangible effect of symbolic matching is that of more robust classification with respect to extraneous responses due to orientation and object configuration. We consider here generic extraneous responses which could be due to unmodeled object features such as additional stores on board the aircraft and extraneous responses due to small errors in the aspect estimate. We show that these problems can be alleviated by eliminating these extraneous responses with nil-mapping.

A less tangible result of the use of a symbolic classifier is the fact that the resulting match of an observed structural description with a candidate

structural description from a library yields an *interpretation* of the observation. This, of course, supposes that the library of structural descriptions have, themselves, meaningful interpretations.

The parametric estimation techniques described in Chapter 2 are used to provide pattern segmentation. The model upon which the estimation procedure is based suggests a distance measure for structural descriptions.

4.1 Parametric Estimation as a Segmentation Procedure

Energy localization implies that individual poles correspond to the objects "scattering centers" or "scatterers". As such, the estimation procedure provides a *segmentation* of the radar measurement. While there exists the possibility of significant *overlap* of the energy in an individual pole with energy contained in other poles of the parametric decomposition in the range dimension, the estimation procedure provides a way of partitioning the energy in the measurement vector into discrete entities which are approximately disjoint in the range dimension.

The following definitions are made for a structural description derived from Prony-based parametric estimates of target signatures. Suppose that the set $Q = \{(p_i, d_i) \mid 1 \leq i \leq M\}$ is a M^{th} order Prony-based decomposition for some radar measurement vector $\vec{\Upsilon}$. The corresponding structural description (see Chapter 3) is $\mathbf{D_Q} = (\mathbf{P_Q}, \mathbf{R_Q})$. The basic idea of deriving a structural description from the Prony-based parametric decomposition is that each node of the structural description consists of a mode (a (d_i, p_i)

pair) minus the angle of the p_i parameter and that this angle parameter determines the relations between the nodes.

Specifically the set of node attributes (T) consists of pole modulus (absolute value) and pole residue, the range of these attributes (the "value sets", V_1, V_2) follows from their definition of these p_i and d_i parameters as

$$T = \{\text{pole modulus, pole residue}\} \tag{4.1}$$

$$V_1 = \mathbf{R}^+(\text{the positive reals}) \tag{4.2}$$

$$V_2 = \mathbf{C}(\text{the complex numbers}) \tag{4.3}$$

then for any node $s_i \in \mathbf{P_Q}$

$$s_i = \{(mode\ residue, d_i), (pole\ modulus, \mid p_i \mid)\}. \tag{4.4}$$

In this case there is only one relation in the the set $\mathbf{R_Q}$. $\mathbf{R_Q} = \{r_1\}$

$$r_1 = (\text{location}, \mathbf{P_Q} \times \mathbf{P_Q}, f_{r_1}). \tag{4.5}$$

The attribute for relation r_1 is the relative *location*,

$$f_{r_1}[(s_i, s_k)] = \frac{-\arg(p_i) + \arg(p_k)}{2\pi} R \tag{4.6}$$

where R is the unambiguous range interval. Note that this definition of the relation r_1 meets requirements 1 and 2 from Chapter 3.

4.2 Distance Measures for Prony-Based Structural Descriptions

If the radar measurement series can be modeled as a series of random variables with known statistics then optimal[1] classification of radar objects consists of choosing that member of a set of candidate measurement series (the library) which minimizes the distance between itself and the observed measurement series. The distance function which provides optimal classification is well known for the assumed white Gaussian noise environment. This distance function is extended to the segmented representation under the assumption that the corrupting noise environment is preserved by the transformation to the segmented representation. That is, noise perturbations to individual elements of the decomposition are probabilistically identical to noise perturbations of the measurement series. This, however, is not the case.

As such, this distance measure between two structural descriptions is sub-optimal except for the special case when it *is* the optimal distance function between the underlying measurement vectors (that case being under the assumption of the trivial correspondence, all poles of one structural description mapped to all poles of the other structural description and no modeling error). It will be demonstrated in Section 4.5, however, how a classifier based on the inter-structural-description distance measure does realize benefits in terms of more robust classification with respect to the optimal classifier.

[1] in a maximum likelihood sense

Since elements of the measurement vector represent frequency domain values, a shift in the range domain of r meters applied to the target corresponds to a phase shift of $-2\pi j f_k(\frac{2r}{c})$ radian being placed on the kth element (we adopt the convention that a positive shift in range corresponds to a shift of the target away from the radar system). Thus, for a given target the measurement vector is a function of the range shift from the point at which the measurements were nominally taken, *viz.*

$$\vec{\Upsilon}(r) = \begin{bmatrix} y_0 e^{-2\pi j f_0(\frac{2r}{c})} \\ y_1 e^{-2\pi j f_1(\frac{2r}{c})} \\ \vdots \\ y_{N-1} e^{-2\pi j f_{N-1}(\frac{2r}{c})} \end{bmatrix}. \tag{4.7}$$

Suppose $C = \{\vec{W}_h \mid 1 \leq h \leq N_c\}$ is a set of complex valued measurement vectors (the library in vector form instead of series form) taken by the given radar system. Each element in this list is from some known object at some nominal range. Suppose that $\vec{O}$ is a measurement vector from some unknown object. We assume that the unknown measurement vector is a noise corrupted version of one of the elements of the library, C, which has been shifted in range by some arbitrary amount, $\tilde{r}$

$$\vec{O} = \vec{W}_h(\tilde{r}) + \vec{G} \quad 0 \leq \tilde{r} < R. \tag{4.8}$$

where $\vec{G}$ is a normally distributed random vector with mean 0 and covariance $\sigma^2 I$. The classification problem is then to determine from which of the library elements the observed unknown measurement vector $\vec{O}$ is derived. Since the noise environment is assumed to be additive Gaussian noise the maximum likelihood solution to this problem is [6]

$$\hat{h} = \overset{\arg\min}{h} \left[\overset{\min}{r} \left\{ \left| \vec{O} - \vec{W}_h(r) \right|^2, 0 \leq r \leq R,\ 1 \leq h \leq N_c \right\} \right]. \tag{4.9}$$

That is, the solution is the measurement vector which minimizes the square distance between it and the observed vector, over all possible range shifts.

Let Ω be a parametric decomposition derived from the observed measurement vector, $\vec{O}$. Let $\vec{o}$ be the vector form of a mode from the decomposition Ω. Then

$$\vec{O} = \sum_{\vec{o} \in \Omega} \vec{o}. \tag{4.10}$$

Similarly, suppose that each element of the catalog $\vec{W}_h$ has an associated parametric decomposition, Λ. The vector form of an element of Λ_h is defined as $\vec{w}_h$

The effect of a shift in range on a measurement vector is reflected in the vector form of the elements of the resulting parametric decomposition. Thus supposing that $\vec{w}_h$ is corresponds to mode $(d, p) \in \Lambda_h$ then the vector form of this mode, under the assumption of a range shift r from the nominal location, is

$$\vec{w}_h(r) = d e^{-2\pi j f_0 (\frac{2r}{c})} \begin{bmatrix} (p e^{-2\pi j \Delta f (\frac{2r}{c})})^0 \\ (p e^{-2\pi j \Delta f (\frac{2r}{c})})^1 \\ \vdots \\ (p e^{-2\pi j \Delta f (\frac{2r}{c})})^{N-1} \end{bmatrix}. \tag{4.11}$$

As is the case when considering full measurement vectors, the distance between a set $A \subseteq \Omega$ and a set $B \subseteq \Lambda$ is defined to be the square distance between sums of the vector form of elements of A and B. Furthermore, this distance reduces to a scalar form using the finite sum formula:

$$
\begin{aligned}
&\left| \sum_{\vec{o} \in A} \vec{o} - \sum_{w(\vec{r}) \in B} w(\vec{r}) \right|^2 \\
&= \sum_{\vec{o} \in A} |\vec{o}|^2 + \sum_{w(\vec{r}) \in B} \left| w(\vec{r}) \right|^2 \\
&\quad + 2\Re \left[\sum_{\substack{\vec{o}_i \in A \\ \vec{o}_k \in A, k>i}} \vec{o}_i^{\,H} \vec{o}_k + \sum_{\substack{w(\vec{r})_i \in B \\ w(\vec{r})_k \in B, k>i}} w(\vec{r})_i^{\,H} w(\vec{r})_k - \sum_{\substack{\vec{o} \in A \\ w(\vec{r}) \in B}} \vec{o}^{\,H} w(\vec{r}) \right] \\
&= \sum_{(d',p') \in A} |d'|^2 \frac{1 - |p'|^{2N}}{1 - |p'|^2} + \sum_{(d,p) \in B} |d|^2 \frac{1 - |p|^{2N}}{1 - |p|^2} \\
&\quad + 2\Re \left[\sum_{\substack{(d'_i,p'_i) \in A \\ (d'_k,p'_k) \in A, k>i}} d'^{*}_i d'_k \frac{1 - (p'^{*}_i p'_k)^N}{1 - p'^{*}_i p'_k} \right. \\
&\quad + \sum_{\substack{(d_i,p_i) \in B \\ (d_k,p_k) \in B, k>i}} d_i^{*} d_k \frac{1 - (p_i^{*} p_k)^N}{1 - p_i^{*} p_k} \\
&\quad \left. - \sum_{\substack{(d',p') \in A \\ (d,p) \in B}} d'^{*} d \epsilon^{-2\pi j f_0 (\frac{2r}{c})} \frac{1 - (p'^{*} p \epsilon^{-2\pi j \Delta f (\frac{2r}{c})})^N}{1 - p'^{*} p \epsilon^{-2\pi j \Delta f (\frac{2r}{c})}} \right] .
\end{aligned} \tag{4.12}
$$

This distance measure, between a subset of the parametric decomposition for the observed measurement vector and a subset of a parametric decomposition for a candidate measurement vector is the distance measure which is used for comparing subsets of nodes from the structural description forms of the parametric decompositions. Recalling from Equation (4.4), elements of $\mathbf{P}_{\Omega}$ (the node set of a structural description derived from the parametric decomposition Ω) are simply the (d', p') pairs of Ω minus the

phase of the p' parameter. Similarly, elements of $\mathbf{P}_{\Lambda}$ can be derived from the elements of Λ.

Supposing now that A represents a set of nodes from structural description $\mathbf{D}_{\Omega}$ and that B represents a set of nodes from $\mathbf{D}_{\Lambda}$, the distance between these sets is made explicit in terms of nodal and relational attributes from the two structural descriptions. In the following formula, a node $s' \in \mathbf{P}_{\Omega}$ is assumed to contain the attributes $(|p'|, d')$ and the relational attribute between nodes s'_i and $s'_k \in \mathbf{P}_{\Omega}$ is denoted $f'_{r_1}[(s'_i, s'_k)]$ so that

$$
\begin{aligned}
& d_s(A,B)(r) = \\
& \quad \sum_{s' \in A} |d'|^2 \frac{1-|p'|^{2N}}{1-|p'|^2} + \sum_{s \in B} |d|^2 \frac{1-|p|^{2N}}{1-|p|^2} \\
& \quad + 2\Re \left[\sum_{\substack{s'_i \in A \\ s'_k \in A, k>i}} d'^{*}_i d'_k \frac{1-(|p'_i||p'_k| e^{\frac{2\pi j}{R} f'_{r_1}[(s'_i, s'_k)]})^N}{1-|p'_i||p'_k| e^{\frac{2\pi j}{R} f'_{r_1}[(s'_i, s'_k)]}} \right. \\
& \quad + \sum_{\substack{s_i \in B \\ s_k \in B, k>i}} d_i^{*} d_k \frac{1-(|p_i||p_k| e^{\frac{2\pi j}{R} f_{r_1}[(s_i, s_k)]})^N}{1-|p_i||p_k| e^{\frac{2\pi j}{R} f_{r_1}[(s_i, s_k)]}} \\
& \quad - \sum_{\substack{s' \in A \\ s \in B}} d'^{*} d e^{-2\pi j f_0(\frac{2r}{c})} \times \\
& \quad \left. \frac{1-(|p'||p| e^{2\pi j(\frac{1}{R} f_{r_1}[(s_{\mathrm{Ref}}, s)]) + \Delta f(\frac{2r}{c}) - \frac{1}{R} f'_{r_1}[(s'_{\mathrm{Ref}}, s')]})^N}{1-|p'||p| e^{2\pi j(\frac{1}{R} f_{r_1}[(s_{\mathrm{Ref}}, s)]) + \Delta f(\frac{2r}{c}) - \frac{1}{R} f'_{r_1}[(s'_{\mathrm{Ref}}, s')]}} \right]
\end{aligned}
\tag{4.13}
$$

Similar comments hold for the nodal and relational attributes of $\mathbf{D}_{\Lambda}$. The nodes $s'_{\mathrm{Ref}} \in \mathbf{P}_{\Omega}$ and $s_{\mathrm{Ref}} \in \mathbf{P}_{\Lambda}$ are reference nodes as were discussed in Chapter 3. The registration parameter r designates the offset in range between the two structural descriptions with respect to these two reference

nodes. In Equation (4.13) the reference nodes are implicitly taken as the "phase center" for both observed and cataloged decompositions.

4.3 Estimation of the Optimal Registration Parameter

In this section, issues involving the computation of the registration parameter r which minimizes the inter-structural-description distance are addressed. The computation of this parameter implicitly calculates the optimal range alignment of the two targets which are represented by the two structural descriptions.

A physical justification for computation of the optimal range offset is that a radar target could be observed at any location within an unambiguous range interval. Therefore, the measurements must be *aligned* in the range domain with each element from the library of targets to account for such possible offsets. A complicating feature of the structural classifier is that the value of the range alignment is a function of the assumed match between the nodes of the two structural descriptions.

In order to minimize the inter-structural-description distance, a modification of the Newton-Raphson method [38] is implemented which is designed to search for minima of the given function. The basic iteration for searching for a minimizing value of a function, $f(r)$ is

$$r_{i+1} = r_i - \frac{f'}{|f''|}. \tag{4.14}$$

where f' indicates a first derivative with respect to the independent variable r and f'' indicates the second derivative. This iteration is applied to the inter-structural-description distance function until adequate convergence is achieved. For our purposes, this is defined as the estimate of the optimizing value of the range offset changing less than a centimeter, out of an unambiguous range interval of 75 meters, between iterations.

Recall that the distance between two structural descriptions is a weighted sum of inter-node-set distances (Equation (3.17)). Thus for a fixed correspondence, $\mathcal{P}$, the required derivatives for the inter-structural-description distance are weighted sums of the first and second derivatives of the inter-node-set distances implied by the correspondence. Closed form expressions for these derivatives are readily calculated from Equation (4.13).

The inter-structural-description distance function can have several minima for the given number of radar measurements. Thus, convergence of the Newton method to a local minimum is a problem. To avoid this problem, the location of each of the local minima must be determined and the least of these used for the optimizing value of the range offset. Since the structural descriptions have associated with them parametric models with range domain interpretations, the approximate location of all local minima of the inter-structural-description distance function is known.

If A and B are singleton nodes of the structural descriptions $\mathbf{P}_\Omega$ and $\mathbf{P}_\Lambda$ respectively then the energy in $A \in \mathbf{P}_\Omega$ and then energy in $B \in \mathbf{P}_\Lambda$ is centered around their respective range locations $f'_{r_1}[(s'_{\text{Ref}}, A)]$ and $f_{r_1}[(s_{\text{Ref}}, B)]$ where s_{Ref} is the reference node in $\mathbf{P}_\Lambda$ and s'_{Ref} is the reference node in $\mathbf{P}_\Omega$.

Thus a minimum of the inter-node-set distance function, $d_s(A, B)(r)$, occurs when the range location of B is co-located with the range location of A; $r = f'_{r_1}[(s'_{\mathrm{Ref}}, s')] - f_{r_1}[(s_{\mathrm{Ref}}, s)]$.

Therefore, the set of approximate locations for the local minima for $d_s(A, B)$ is the set of *differences* in the range locations of A with respect to the range locations of B, $\{r_{a,b} \mid r_{a,b} = f'_{r_1}[(s'_{\mathrm{Ref}}, s')] - f_{r_1}[(s_{\mathrm{Ref}}, s)],\ s' \in \mathbf{P}_\Omega s \in \mathbf{P}_\Lambda\}$. An arbitrary inter-structural-description distance function is a sum of inter-node-set distance functions. Therefore, the set of approximate local minima used for initialization of the Newton optimization performed on the inter-structural-description distance function is the union of the sets of approximate locations for the local minima of the constituent inter-node-set distance functions.

Note that the band-limited nature of measurement response effectively induces a *modulation* on the inter-node-set distance functions and thus on the inter-structural-description distance functions. Under a narrow-band assumption this modulation can be effectively removed by using the complex envelope of the distance function. Thus the inter-structural-description distance function which is optimized separates out the range-dependent terms and takes twice their absolute value instead of twice the real part as is indicated in Equation (4.13).

4.4 Distance Function Based Relational Constraints

The criterion for determining if a particular pairing of two elements in the candidate partition is relationally consistent with the input partition is defined in terms of "half-distance" intervals surrounding the optimizing values of the registration parameter (range offset) for both the candidate pairing and the input partition.

The *half-distance* interval is the interval surrounding the optimizing values of the registration parameter in which the distance does not change by more than a factor of two from the optimized value. This definition serves for the distance function associated with both a candidate pairing and the input partition. In the case of a candidate pairing, the distance function used for determining the half-distance interval is the inter-node-set distance between the two elements of the candidate pairing. In the case of an input partition, the distance function used for determining the half-distance interval is the implied inter-structural-description distance. Since the only registration parameter which is currently used is an offset in range, these intervals will be referred to as the *half-distance rangewidths*.

The definition of relational consistency which has been adopted for the current experiments is the following: If the optimizing value for a candidate pairing lies within the half-distance rangewidth of the input partition and the optimizing value for the input partition lies within the half-distance rangewidth of the candidate pairing then the pairing is said to be relationally consistent with the partition.

Recall from the previous section that optimization of a candidate match over the registration parameter is accomplished numerically. A by-product of this minimization is the first and second derivatives of the distance function with respect to r. By approximating the distance function with a quadratic function at the optimized value of r, the half-distance interval can be accurately approximated with the values of the derivatives of the distance functions.

In practice, this criterion produces a significant reduction in the number of partitions for which optimization of the registration parameters is necessary. For the experiments described in this study the maximum number of nodes in any given partition element fixed at three. Under this restriction, the number of possible partitions as a function of the total number of nodes, is given in Table 4.1. Also included are typical values for the number of partitions which are actually evaluated for optimum range offset in a simulation which uses the relational consistency criterion. This criterion appears to produce a reduction of 75% to 95% in the number of partitions actually evaluated.

Table 4.1: Number of possible partitions and typical numbers evaluated.

Total nodes, candidate and prototype	Distinct partitions	Typical number of partitions
6	166	40
8	2780	300
10	61136	3200

4.5 Experimental Results

The results of simulation studies of the matching algorithm are reported in this section. Given here are estimates of the percent misclassification of the matcher for comparison to the standard statistical classifier as well as descriptions of the resulting correspondences.

Descriptions of the matches resulting from minimization of an inter-structural-description distance are provided to outline the benefit derived from a structural approach to object recognition. We will demonstrate how such a match can provide a description of an object and how increased robustness comes about from a structural approach.

The experiments described here use data obtained from the compact radar range at The Ohio State University. The range and measurement process is fully described in [32]. Scale models of five commercial aircraft are used. The resulting measurements are scaled accordingly. The range of frequencies which were used in these experiments correspond to 30 to 60 MHz in scaled frequency.

Figure 4.1 depicts the down-range signature envelope of an object as defined by Equation (2.12) Also indicated on this figure are correspondences between a set of measured scattering centers and a set of scattering centers from a particular library element, as determined by minimization of the inter-structural-description distance. In this figure, scatterers defined by the parametric segmentation procedure are indicated symbolically on the envelope impulse response for both library and unknown. Scatterers which are from the unknown or measured object are labeled $A, B, C, \ldots$

while those from the library object are labeled $1, 2, 3, \ldots$. Correspondences between measured and library scatterers are indicated by the shaded regions. For example, in Figure 4.1 scatterer "A" from the library object is matched to scatterer "1" from the observed object.

The correspondences depicted in Figure 4.1 are determined by minimizing an inter-structural-description distance function with a link cardinality factor, α, of 1. The indicated range offset parameter corresponds to a 1.8540 meter shift of the library object to the "right".

The resulting correspondences illustrate a number of properties of the matching algorithm. For example, the algorithm tends to match similar scatterers, as in the match between scatterers "A" and "1". In addition the algorithm indicates the absence of a corresponding scatterer when none is present as in the indicated lack of a match for the combination of scatterers "D" and "E". Furthermore, the resulting correspondence between the scatterers yields an "interpretation" of the unknown object with respect to the prototype.

In the current example, the resulting match can be interpreted as indicating that the observed object resembles the prototype object (with the resulting value for the inter-structural-description distance indicating the relative likeness) with an extra scatterer on one end and an extra scatterer in the middle. The actual envelope impulse responses given in Figure 4.1 correspond to the same physical object at different aspects angles (40° and 45° "yaw" with respect to the radar system). Conceivably, such an interpretation could serve as a hypothesis to be combined with other information and tested by a higher level reasoning process.

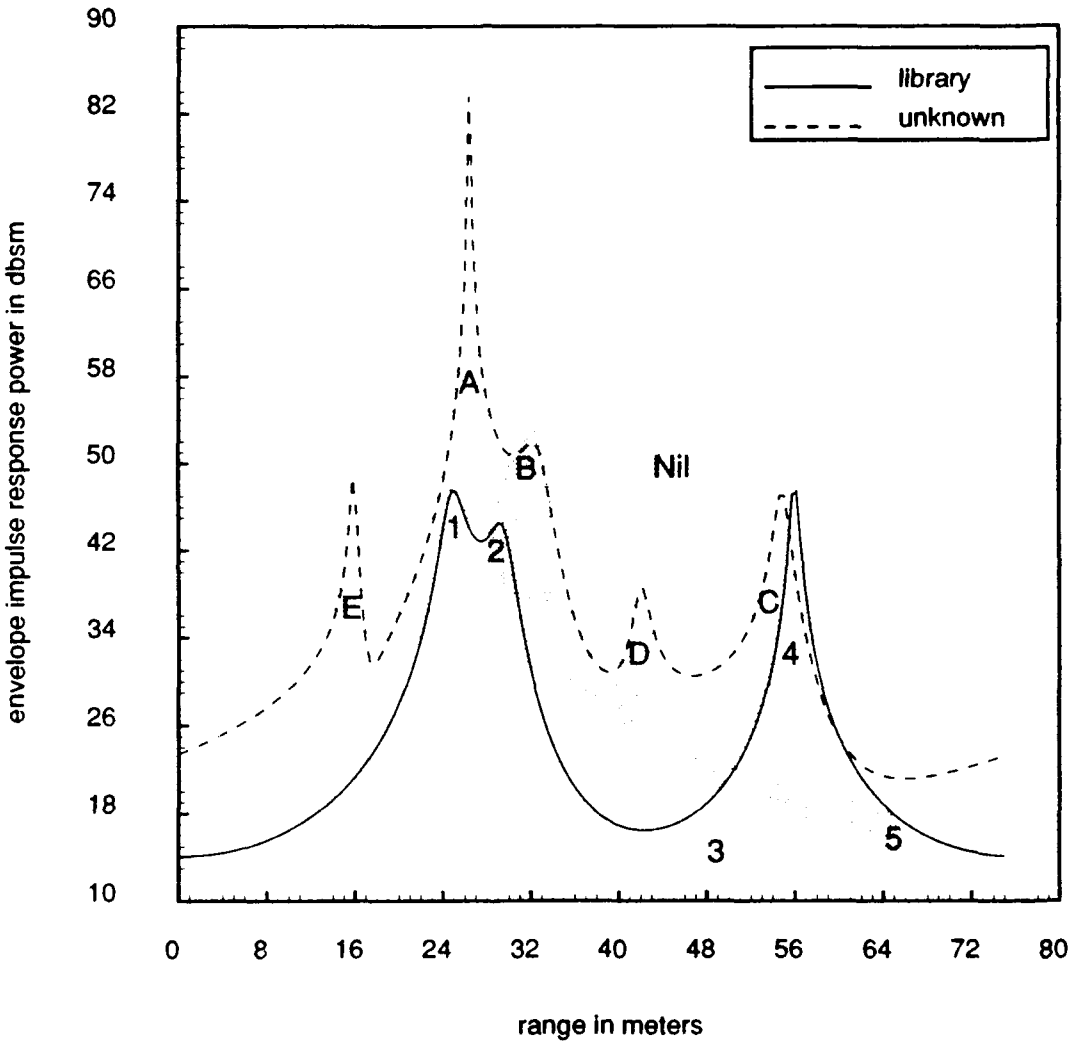


Figure 4.1: Correspondence between poles of prototype (library) and measured objects. Prototype and measured objects are same object at slightly different aspects.

The energy, range and pole magnitude (modulus) of the scatterers depicted in Figure 4.1 are given in Table 4.2. In addition, the correspondences between the poles which minimize the inter-structural-description distance are indicated by the individual rows of this table. The numbers in the "Range" column indicates where the center of the energy for the particular scatterer lies with respect to an arbitrary reference point. The numbers in the "Energy" column represent the total energy contained in the scattering center, integrated over the entire unambiguous range interval. The "Modulus" parameter is the magnitude of the p parameter of a scatterer. As discussed in subsection 2.3.1, this parameter gives an indication of the energy concentration of the mode in the range dimension. Values of modulus close to 1 indicate "impulsive" scatterers, those away from 1 (greater than 1.1 or less than .8) indicate a more "distributed" scatterer.

Table 4.2: Pole parameters for correspondence example.

Library, prototypical scatterers				Observed, unknown scatterers			
Scatterer	Range	Energy	Modulus	Scatterer	Range	Energy	Modulus
1	25.1033	989	1.0867	A	26.6696	15554	1.0059
2	29.7519	845	1.0960	B	32.8855	2039	0.9016
5	69.1534	30	1.9469				
3	50.6451	58	1.5259	C	54.9817	275	0.9451
4	56.0781	44	1.0317				
Nil				D	42.1804	163	1.0649
				E	16.2270	86	0.9764

Comparison of the parameters between matched sets of scatterers illustrates a number of properties of the minimum distance matcher. In

particular, this table verifies the previous assertion that similar scatterers are matched together. Scatterer "A" is the highest energy scatterer from the prototype object and scatterer "1" is the highest energy scatterer from the observed object. In addition, the two scatterers have similar values for their modulus.

The match made between scatterers "2" and "5" of the prototype object and scatterer "B" of the unknown object points out another interesting aspect of the matcher. The match between scatterers "2" and "B" is intuitively obvious. However, inclusion of scatterer "5" in the match seems counter intuitive due to its location, seemingly far from the involved scatterers. Note that scatterer "5" is of relatively high modulus, this implies that the energy in this scatterer is distributed almost uniformly over the entire unambiguous range interval. As such, the indicated location for scatterer "5" is almost meaningless. Coupling scatterer "5" with scatterer "2" only serves to better match the energy in scatterer "2" to that of scatterer "B". Similar comments hold for the matching of combination of scatterers "3" and "4" to scatterer "C".

Monte-Carlo simulation is used to produce performance estimates for several forms of the structural based classifier as well as for a maximum likelihood classifier which operates directly on the measurement vector for comparison purposes. In addition performance estimates of a classifier which makes the trivial correspondence (all poles of one structural description mapped to all poles of the other structural description) are made to estimate the effect of modeling error.

The results of this simulation are given in Figure 4.2 and Figure 4.4. In both figures, the estimated probability of misclassification as a function of the added noise power (in decibels above 1 square meter or dbsm) is given for each of the classifiers. The solid curve, marked "Range Aligned M.L." is the performance estimate for the statistically motivated maximum likelihood classifier. The curve marked "Post-proc N.N." (post processing nearest neighbor) corresponds to the "trivial correspondence" classifier described above. The other curve corresponds to the structural classifier with the value of the link cardinality factor, α, equal to 1 and the nil-map propensity, ϕ, equal to 0.5. Thus, this particular classifier encourages both "linking" and "nil-mapping".

The results depicted in Figure 4.2 are from simulations in which the elements of the library of radar measurement vectors corresponds to a set of 5 commercial aircraft, all at "nose-on" aspect with respect to the radar system. The unknown measurement vectors presented to the simulated classifier are exactly those from the library with white Gaussian noise superimposed on the measurements. The average "signal levels" of the individual library elements vary from 12 to 23 dbsm with a library average of approximately 19.5 dbsm.

Since all 5 library objects are presented to the classifier with equal probability (assumption of equal priors) a classifier which chooses one of the library classes without any information would achieve a misclassification probability of .8. Thus any overall misclassification probability beyond .8 represents a misuse of the information presented to the classifier. None of

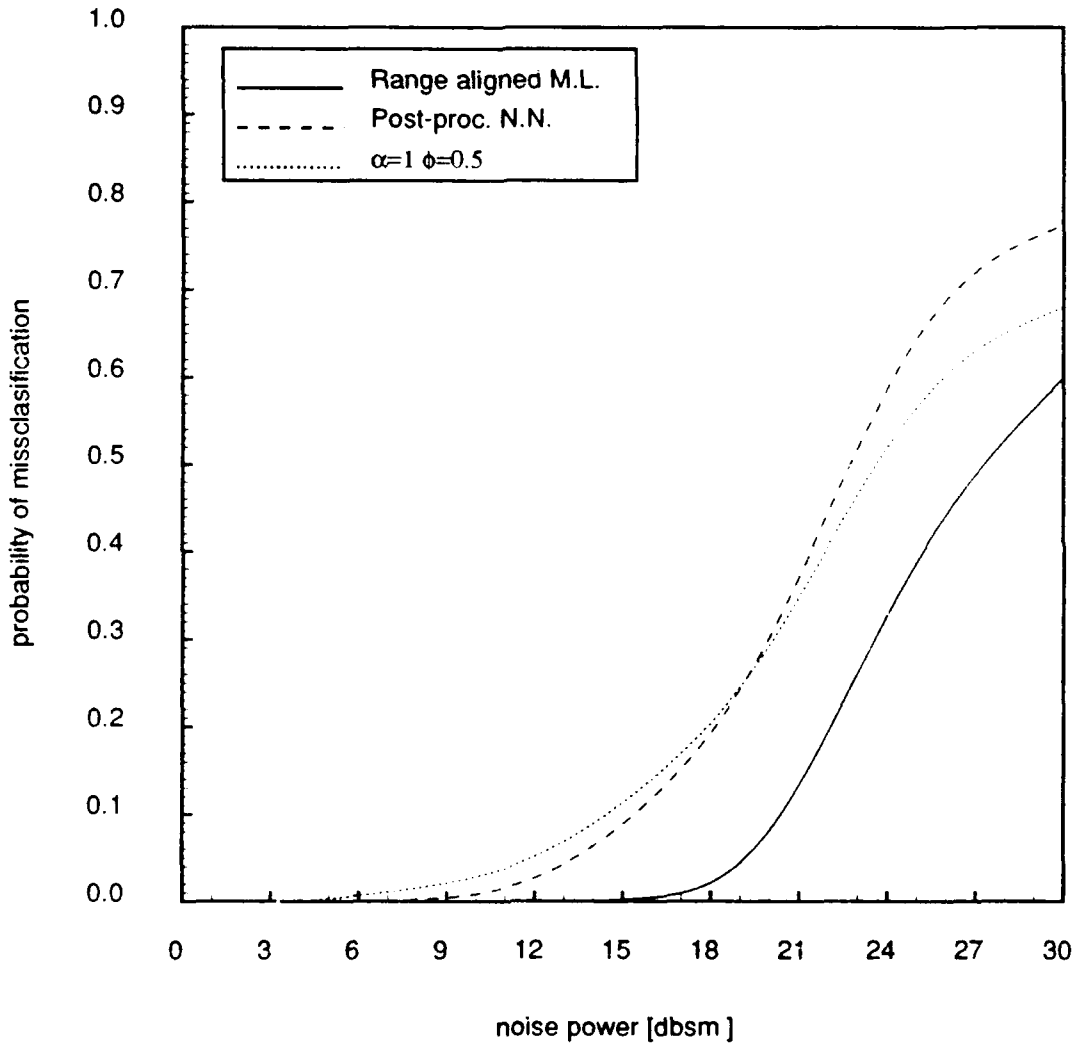


Figure 4.2: Classifier Performance with "nose-on" library and set of unknowns.

the classifiers have produced performance estimates beyond .8 for the cases tested.

Comparison of the misclassification results for the structural based classifier to that of the maximum likelihood decision rule indicates that the structural classifier is able to classify nearly as well as the statistically optimal decision; the performance of this classifier lagging the optimal classifiers by approximately 6 dbsm at the 10% misclassification level. Similar results are obtained for structural classifiers with different values of the link cardinality factor and nil-map propensity.

The reason for the performance compromise associated with the structural classifier can be at least partially explained by the modeling error due to reduced data dimensionality of the parametric representation. The simulated measurement vectors for the current case consist of 16 complex scalars ($N = 16$). The model order used for the parametric estimation step is 3. Thus, this step represents almost a 3:1 reduction in data dimensionality. In addition, the detrimental effects of the parametric estimation step is indicated by the performance estimate of the "Post-processing nearest neighbor" classifier. Recall that this classifier corresponds to a "trivial correspondence" structural classifier and as such is a statistically optimal decision rule when there is no modeling error. The post-processing nearest neighbor classifier however yields performance almost identical to that of the structural classifier, and for lower signal to noise ratios, the indicated performance is inferior.

A case in which the structural classifier exhibits superior robustness with respect to the "maximum likelihood" classifier is now described. In this case the unknown objects presented to the classifier have an "extraneous" scatterer appended and a "true" scatterer deleted from each object (in addition to the the unknown measurements being corrupted by white Gaussian noise). Such an addition and deletion is illustrated in Figure 4.3 (without any noise corruption). In this figure, the envelope impulse response of a library element is depicted with its scatterers labeled (1,2,3,4) as in Figure 4.1. The corresponding unknown pattern which is derived from this library element (for presentation to the simulated classifier) is also given here with its scatterers labeled (A,B,C,D). Note that the unknown envelope response resembles the library response from which it was derived with two important exceptions. These being the large scatterer on the "front" of the object and the deletion of the small scatterer from the "middle" of the object. For the Monte-Carlo simulation, similar distortions were made to each of the library objects to generate a set of unknown representations to present to the classifier.

Figure 4.4 gives the performance estimates for the three classifiers described above with "extraneous scatterer" unknowns. Note that for sufficiently low noise power, the structural-based classifier exhibits superior performance to the maximum likelihood classifier. This is because the maximum likelihood classifier misclassifies the object depicted in Figure 4.3 in a noiseless environment while the structural classifier does not. Furthermore, the structural classifier makes the correspondences indicated in Figure 4.3, correctly "nilling" the extraneous scatterer in the unknown and indicating

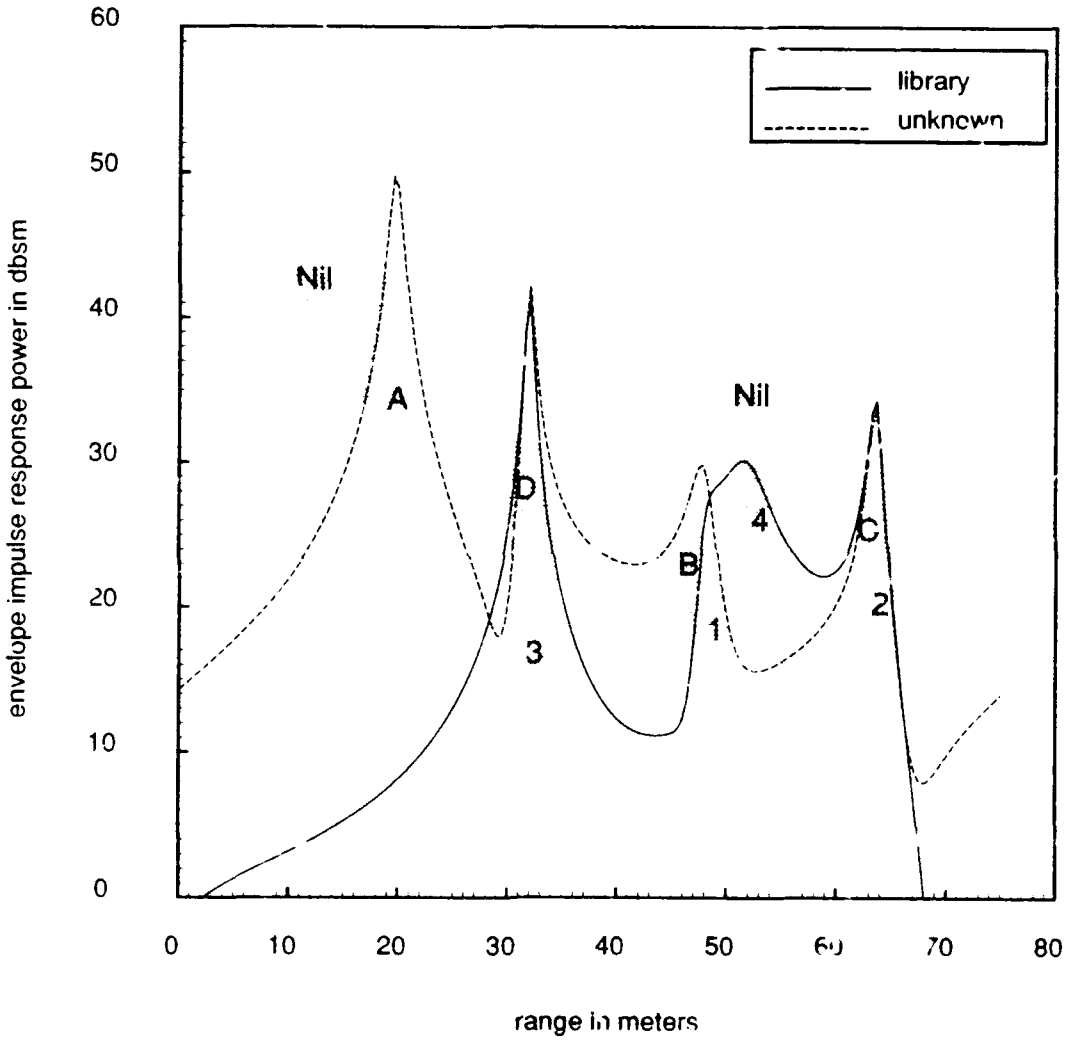


Figure 4.5: Correspondence between scatterers of prototype (library) and measured objects. Measured object is the prototype object with one extraneous scatterer appended and one deleted.

the deleted scatterer in the library object and correctly linking the remaining scatterers. The parameters for the constituent scatterers of Figure 4.3 are given in Table 4.3.

Table 4.3: Pole parameters extraneous scatterer example.

Library, prototypical scatterers				Observed, unknown scatterers			
Scatterer	Range	Energy	Modulus	Scatterer	Range	Energy	Modulus
Nil				A	20.0000	2000	1.0500
1	48.1140	157	1.0916	B	48.1140	157	1.0916
2	63.7117	71	1.0464	C	63.7117	71	1.0464
3	32.1584	42	0.9682	D	32.1584	42	0.9682
4	51.8176	29	0.8200	Nil			

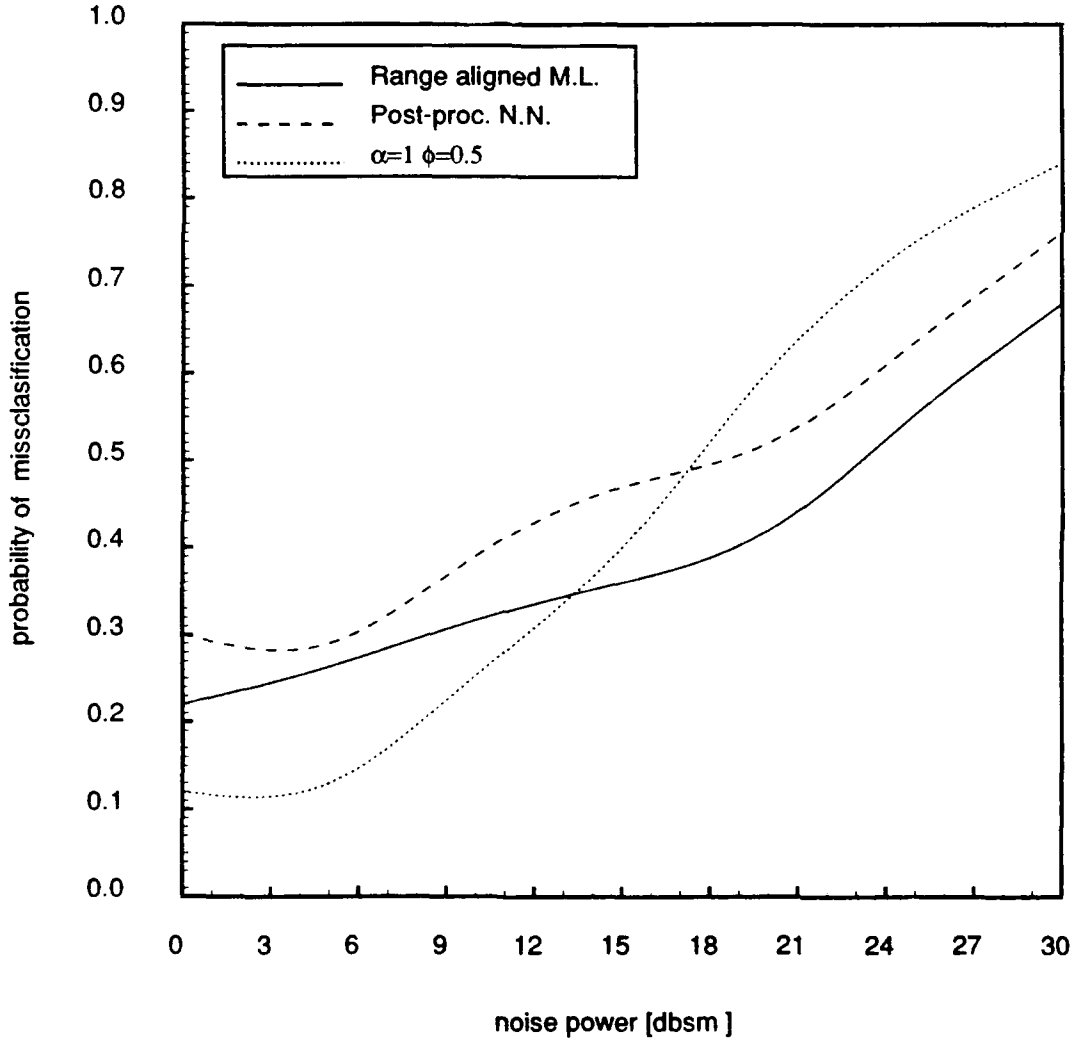


Figure 4.4: Classifier performance with 1 added extraneous scatterer.

Chapter 5

CONCLUSIONS

5.1 Summary

In this report, a framework for a pattern recognition methodology has been presented which includes a number of extensions to the theory regarding matching of symbolic representations. These extensions have been made to accommodate the statistical nature of Prony-based parametric decompositions of radar measurement vectors. The result is a pattern recognition system which classifies unknown patterns in a way which is intimately related to a statistically optimal classifier and which also provides a description of the unknown target in the context of the given library of possible targets.

Prony-based parametric estimation, as applied to a low frequency radar measurement series, has been chosen to provide a segmentation of the radar measurement series since it exhibits many desirable properties (including increased resolution, data reduction etc.) for representation the features of the radar target. However, under noise perturbations, the random nature of a parametric decomposition is yet fully characterized. Thus, a method by

which statistically optimal categorization of an unknown radar target from an observed parametric decomposition cannot be immediately determined from the specification of the extraction algorithm and measurement process.

The discrete nature of a parametric decomposition suggested the use of a symbolic matching procedure. However, the theory for matching of symbolic representations is inadequate for the matching of elements of parametric decompositions for a number of reasons.

Noise perturbations, in the radar measurement series cause not only variations in the constituent parameters of a parametric decomposition but also affect the number of resultant modes and the relationship between noisy modes and noiseless modes. For example, there exist circumstances in which noise perturbations can cause modes to "split", "merge" or "disappear". This can occur even when the number of estimated modes is fixed for both the noiseless and noisy modes.

In addition, prior theories pertaining to matching of structural descriptions do not consider the importance of matching in a *symmetric* sense. Non-symmetric matching is applicable to circumstances in which the recognition of a sub-patterns within a given pattern is desirable. In these cases there is no penalty associated with "nil-mapping" elements of one of the structural descriptions involved in the match. The use of non-symmetric matching for the radar problem has implications which are undesirable. For example, with non-symmetric matching it is possible for the matcher to find a very low energy target within an extremely high energy target, thus declaring the high energy target the optimal decision.

Furthermore, probabilistic models of structural descriptions have included no provision for accounting for statistical dependencies among the constituent elements of the symbolic representation. The nature of the Prony estimation process indicates that statistical dependencies are an integral part of a valid statistical modeling of the estimation process.

The major contribution of this report is the development of a pattern recognition system which addresses all of the above mentioned problems with application of structural description matching to radar target identification. This has been accomplished with two modifications to the symbolic matching formalism. The first of these is a reformulation of the matching task in terms of "partitions" of the elements of the two symbolic representations instead of a mapping of the elements of one of the symbolic representations to the elements of the other symbolic representation. The second modification is the inclusion of "registration parameters" to describe the relational aspects of the symbolic representations.

It has been demonstrated how the use of the partition formalism brings a natural sense of symmetry to the resultant matches as well as allowing for a completely general match, i.e., matches between elements of the two symbolic representations are not restricted to be 1-1 or 1-$\emptyset$, etc. This allows for modeling of "merging" and "splitting" of elements of the parametric decompositions as described above. Thus the resulting pattern recognition technique realizes the capability of autonomously determining the sense of the match between elements of the symbolic representations which, in turn, allows for more accurate object interpretation.

We have also demonstrated how the use of registration parameters in the matching of structural descriptions allows for an implicit modeling of statistical dependencies among elements of the parametric decompositions. In addition it has been shown how these parameters allow for deformation of relational attributes to the same degree as conventional relational models, while more fully making use of the available contextual information. Under the registration parameter formalism, the relational portion of the symbolic representation may still be used to reduce the search space as is done in "relational consistency" checks for structural description matching.

In order to help identify those circumstances in which the given framework for matching is applicable, we have made explicit a number of qualifying assumptions regarding the relational portion of symbolic representations and the inter-symbolic-representation distance function. We see the specialized structure imposed upon the symbolic representation as natural and unrestrictive in light of current applications. The qualification have been shown to be sufficient for the synthesis of a pattern recognition scheme in which a complete, synergetic use of semantic and structural information is realized.

The "partition" formalism, developed here, has allowed for the explicit proof that the metric property is not desirable for inter-node-set distance function. This is due to the fact that the metric property eliminates the possibility of all but a trivial correspondence between symbolic representations. Conversely, if a statistical model of the problem at hand indicates a metric inter-node-set distance then matching of discrete entities should not be considered.

Finally, we have demonstrated how the formulated matching theory produces a classifier with performance rivalling a statistically optimal classifier under standard assumptions of known signal parameters and noise environment. We have also demonstrated that the developed classifier produces performance which surpasses the optimal statistical classifier in less benign circumstances. Furthermore, since the matching process is symmetric in nature, the resulting correspondences between parametric decompositions allow an interpretation of the observed pattern representation in which statements can be made regarding both what is missing in addition to what is extraneous, in a consistent fashion.

5.2 Concluding Remarks

The "naturalist" viewpoint adopted here has lead to a number of conclusions. We offer the following comments as insight gained into the structural forms of pattern recognition applied to symbolic representations derived from sensor data as well as the target identification problem.

Whenever possible, the distance measure between elements of the symbolic representations should be directly induced by the statistical description of the pattern generation process. Indeed, the entire symbolic matching strategy presented here is driven by the statistical and physical description of the pattern generation process. In situations where the distance function cannot be deduced from knowledge of the pattern generation process, approximate or heuristic formulations may be employed.

In a traditional model for matching of symbolic representations, the segmentation process implicitly discards information regarding the relation of input patterns to "the world". Only relational information regarding the elements of the symbolic representation is retained. Such "absolute" information should be regarded as necessary but arbitrary; when it is unavailable it is manufactured in the form of a synthesized relations to an artificial reference node. Removal of such information unnecessarily complicates the classification task. Thus, for classification and description of sensor-based symbolic pattern representations, the primary descriptive power of structural descriptions comes from the discrete nature of the representation.

By considering matching in the more general sense, the matching algorithm is able to account for certain types of errors in the segmentation process. Even in cases in which the more generalized sense of matching is not indicated by a modeling of the pattern generation process, the use of a simpler, possibly error-prone segmentation scheme may allow for a probabilistic modeling of the symbolic representations. In such a case, errors in the segmentation scheme would be allowed for in the matching process.

The use of a limited number of measurements implies that the result of the segmentation process contains a limited number of scattering centers. In this way, the number of radar measurements dictates the resulting "range resolution". The time required for the matching of two symbolic representations grows rapidly with the total number of unknown and catalogued scattering centers. Thus, by using only a limited number of measurements, matching of symbolic representations, as described here, becomes practical. This indicates that the algorithm is most applicable to matching of pattern

representations which have limited resolution, at least until execution of the algorithm can be accomplished in a significantly shorter period of time.

5.3 Additional Notes

An apparent solution to the problem of matching of structural descriptions with large numbers of discrete symbols is to use a hierarchic symbolic representation and matching scheme [39]. By hierarchic we mean that the representation exists on multiple levels of resolution simultaneously. In this sense, each element of a symbolic representation on a fixed level of resolution can be "broken up" into several, more finely resolved elements on a "lower level" of resolution (note that by lower level of resolution we mean finer resolution). Under this scheme, a given symbolic element represents, constituent lower level elements in a gross, macroscopic way. The lowest level representation contains the largest number of symbolic elements and each higher level contains fewer total elements than the previous. Thus, each level of resolution represents the same pattern on a different resolution scale, the difference between the levels being the relative coarseness with which the pattern is represented.

A hierarchic symbolic representation may be devised such that the number of symbolic elements at the highest level of resolution is small enough to allow matching in a relatively short period of time. Matching of symbolic representations with a hierarchic structure begins by first matching unknown and catalogued symbolic representations on the coarsest resolution (the "highest" level of the hierarchy). The resulting correspondence is

then used to commence a number of matching procedures on the next lower level of resolution. This process is repeated until correspondences between elements of the two symbolic representations at the lowest level of resolution is obtained. If, for any given element of a symbolic representation, the number of lower level symbols is restricted to be a sufficiently small number then each matching procedure can run in a relatively short period of time. The time required for the overall matching procedure between the lowest level elements is therefore also limited.

If elements of a given hierarchic symbolic representation on all resolution levels are of the same form then use of the generalized sense of matching presented here can be used to directly determine correspondences between symbolic representations of differing resolution levels. For the target identification task, such a feature has immediate use for determining the class of a target in which its representation was derived from fewer than the normal number of measurements. This feature could also be used for investigating the properties of differing segmentation (parametric estimation) schemes as well as the effect of data length on the resulting symbolic representation.

In addition this theory can be directly extended to similar problems in which extraction and categorization of modal elements of a signal is important. Furthermore, there exists the possibility of applying this theory to other areas such as matching the states of two time-reversible Markov chains or other problems in which a symbolic model is natural.

In addition there are a number of areas in which the theory presented here could be expanded or improved. For example, the incorporation of additional registration parameters such as rotation and scaling. The incor-

poration of these additional registration parameters could possibly induce complications in the implementation; optimization of a function over a single variable is a relatively straight-forward procedure while optimization over many variables is not.

Furthermore, there are a number of areas which have not been fully investigated with respect to reduction in the amount of computations necessary for making a match. The current implementation takes a considerable amount of time to execute. Proposed methods for computation reduction include limiting the number of matches which may be considered at a given level of recursion, both in a sub-optimal manner and, when combined with a limit on the number of elements allowed in a particular partition element, in an optimal fashion.

Up until this point, the term non-symmetric matching has meant that nil-mapping of elements of either the candidate or a prototype structural description can occur without contribution to the match quality function. In addition, the inter-structural-description distance function has been required to be a symmetric function. However, it has been suggested that the use of a non-symmetric distance function (which would imply a non-symmetric matching) could have use for noise-cleaning or could possibly find use in incorporation of information regarding presence or absence of certain of a given targets' scattering centers. Thus, the definition of non-symmetric matching is altered and methods of exploiting the non-symmetry should be investigated.

Appendix A

DETERMINATION OF CORRESPONDENCES BETWEEN ELEMENTS OF TWO PARAMETRIC DECOMPOSITIONS

The purpose of this appendix is to provide a formulation of the implementation of the matching algorithms which determine the correspondences between two parametric decompositions. The first section covers a simplified version of the full matching algorithm described in Chapter 3. This algorithm is limited in terms of the type of correspondences it is capable of making between parametric decompositions. It does, however, contain the potential to produce extremely fast correspondences which minimize the given distance function.

The second algorithm described is an implementation of the matching algorithm described in Chapter 3. This algorithm considers matching in an extremely general sense. In its most general form, this algorithm considers

matches between the two parametric decompositions which are not only 1-1 but those which are 1-$\emptyset$, 2-1, 2-$\emptyset$, 3-1, 3-2 and so on. In addition, this algorithm as implemented optimizes the value of the "range alignment" (the maximizing value of the "range" correlation value) between the targets for each correspondence considered. This algorithm thus minimizes the distance over the set of all correspondences and range offsets.

Both matching algorithms are described as a recursion on a matrix. The basic recursive step for each of these algorithms is to generate a number of reduced order (reduced by 1) matrices from the input matrix and pass each of the reduced order matrices to a lower level of recursion. We refer to such matrices which are generated and passed on as being *operational.*

In both of these algorithms, matrix reduction is implicitly equivalent to the incremental formation of a correspondence between the elements of the parametric decompositions. The differences between the algorithms lie in the type of matrix upon which they operate and the way in which a lower level matrix is generated from an upper level matrix.

The elements of the operational matrices consist of terms, or sums of terms, from a formula for the magnitude square distance between the measurement vectors from which the parametric decompositions are derived. Suppose that two measurement vectors, $\vec{\Upsilon}_A$ and $\vec{\Upsilon}_B$ have parametric decompositions, Λ_A and Λ_B respectively. Then, assuming no modeling error, the mean square distance between the two measurement vectors can be expressed in terms of elements of the parametric decompositions in vector form, *viz.*

$$\left|\vec{\Upsilon}_A - \vec{\Upsilon}_B\right|^2 = \left|\sum_{\vec{v}_a \in \Lambda_A} \vec{v}_a - \sum_{\vec{v}_b \in \Lambda_B} \vec{v}_b\right|^2 \tag{A.1}$$

$$\begin{aligned} &= \sum_{\vec{v}_a \in \Lambda_A} |\vec{v}_a|^2 + \sum_{\vec{v}_b \in \Lambda_B} |\vec{v}_b|^2 \\ &\quad + 2\Re\left[\sum_{\vec{v}_a \in \Lambda_A,\ \vec{v}'_a \in \Lambda_A,\ \vec{v}_a \neq \vec{v}'_a} \vec{v}_a^H \vec{v}'_a \right. \\ &\quad + \sum_{\vec{v}_b \in \Lambda_B,\ \vec{v}'_b \in \Lambda_B,\ \vec{v}_b \neq \vec{v}'_b} \vec{v}_b^H \vec{v}'_b \\ &\quad \left. - \sum_{\vec{v}_a \in \Lambda_A,\ \vec{v}_b \in \Lambda_B} \vec{v}_a^H \vec{v}_b \right]. \end{aligned} \tag{A.2}$$

In Equation (A.2) $\vec{v}_a$ and $\vec{v}_b$ represent the vector form of the elements (modes) of the parametric decompositions Λ_A and Λ_B respectively, as in Equation (2.10). Equation (A.2) gives the amount of residual energy associated with matching the measurement vector Υ_A with measurement vector Υ_B. In non-parametric methods of classification, the residual energy is minimized over the set of all catalogued measurement vectors to determine the overall best match for a given observed measurement vector.

We show in this appendix how algorithms which operate on matrices with elements consisting of the terms of Equation (A.2) determine "minimum distance" correspondences between elements of the parametric decompositions, Λ_A and Λ_B. The distance associated with a particular correspondence is equal to the sum of the residual energies associated with each grouping of elements implied by the correspondence. Thus determination of optimal correspondences is defined by minimization of the total residual energy.

For the first algorithm, elements of an operational matrix represent the distance between the discrete elements of two parametric decompositions. Thus, the first algorithm described is termed the *d-matrix algorithm*. For the second algorithm elements of the matrices represent terms of a distance function which describes the distance between two measurement vectors. Elements of this matrix are used to minimize a distance between *sets* of elements from the two parametric decompositions. This algorithm is therefore referred to as the *generalized d-matrix algorithm*.

A.1 The d-matrix Algorithm.

The d-matrix algorithm is applicable to the determination of correspondences between elements of two sets in which the pairwise distances between elements of the sets are available. This algorithm is currently implemented in a "static" sense, i.e., there exists no facility for adjusting the "range" location of the elements of one parametric decomposition with respect to the other. It is therefore applicable to the estimation of the parameter statistics for a given parametric decomposition.

The objective of the d-matrix algorithm is to select, from the set of modes of one of the parametric decompositions, a match for each element of the other parametric decomposition in such a way as to minimize the total residual energy. Another way of viewing this task is in terms of a search for a 1-1 mapping $\mathcal{M} : \Lambda_A \mapsto \Lambda_B$ in which each element (mode) of Λ_A is "canceled" by its corresponding element in Λ_B (or possibly the "nil" mode, i.e. the mode with no energy) to a maximal extent. Thus, the quantity which is minimized is

$$\sum_{\vec{v} \in \Lambda_A} |\vec{v} - \mathcal{M}(\vec{v})|^2 . \qquad (A.3)$$

over the set of all 1-1 mappings. Note that under this scheme all of the energy in the parametric decomposition Λ_A is included in the calculation of match quality. Only those elements of Λ_B which are paired with elements of Λ_A have their energy included in the match quality function. Thus the resulting matcher is non-symmetric. This brings about several questions when the target recognition task is considered. Is such a match desirable? Should the mapping be from the observed parametric description to the library parametric description or vice versa? It is due to these complications that non-symmetric matching is not fully investigated for the target recognition application.

Supposing the modeling equation (Equation (2.9)) holds exactly for both parametric decompositions and the two measurement vectors are equivalent, then the two resulting parametric decompositions are equivalent to within a "labeling" of the modes. Minimization of the quantity given in (A.3) results in the renaming of the modes of one of the parametric decompositions with the "names" of the modes which they are most "like" in the other parametric decomposition in an overall sense.

A.1.1 Operation of the d-matrix Algorithm

The d-matrix algorithm begins by calculating a matrix of pairwise distances or residual energies between elements of two parametric decompositions. Supposing that elements of the parametric descriptions, Λ_A and Λ_B, have

an arbitrary ordering imposed on them and that $|\Lambda_A| = M_a$ and $|\Lambda_B| = M_b$, then

$$\Lambda_A = \{\vec{v}_{a_i} \mid 1 \leq i \leq M_a\} \tag{A.4}$$

$$\Lambda_B = \{\vec{v}_{b_i} \mid 1 \leq i. \leq M_b\} \tag{A.5}$$

The elements of the input value of the d-matrix are the residual energies resulting in matches between the elements of Λ_A and Λ_B. Columns of this initial value for the d-matrix correspond to the elements of Λ_A and the rows correspond to elements of Λ_B or "nils". Sufficient numbers of "nil-mode" rows are appended to allow for nil-mapping of the entire set of modes from Λ_A. Elements of the input d-matrix, denoted $d_{i,j}$, are therefore

$$d_{i,j} = \begin{cases} \left|\vec{v}_{a_j}\right|^2 + |\vec{v}_{b_i}|^2 - 2\Re\left(\vec{v}_{a_j}\vec{v}_{b_i}^H\right) & 1 \leq i \leq M_b \\ \left|\vec{v}_{a_j}\right|^2 & M_b + 1 \leq i \leq 2M_b \end{cases} . \tag{A.6}$$

Elements of the d-matrix also have a simple form in terms of the (p, d) form of the modes of the parametric decompositions.

The minimization algorithm is now described. Input to the algorithm (that is, the subroutine is initially called with these values) are the initial value of the d-matrix, **DMAT**, a parameter for the globally applicable minimum distance achieved, **GLOBDMIN** and a parameter which indicates the accumulated distance for the correspondence which has been accumulated so far, **DSOFAR**. Since, at initialization, there has been no correspondence made the global minimum distance parameter is set to ∞ and the accumulated distance is set to 0.

The output is the 1-1 mapping from Λ_A into Λ_B which minimizes the distance between the two structural descriptions and the distance (residual

energy) associated with this mapping. The mapping, named **PERM**, is a row vector, one element for each column of the d-matrix, which assigns to each column a row. Recall that the columns of the d-matrix have an exact 1-1 correspondence to the elements of Λ_A and the rows of the d-matrix have an exact 1-1 correspondence to elements of Λ_B (or "nils"). As such, choosing an element from the d-matrix is equivalent to forming a match between the elements of Λ_A and Λ_B (or possibly nil). The distance associated with this mapping, **DMIN**, is the sum of the elements in the d-matrix implied by the given mapping.

The algorithm is given in the form of a recursive subroutine. Therefore, all variables are local except the globally applicable minimum distance. This variable, **GLOBDMIN**, must be declared global since it is the value of the minimum distance which has been achieved so far, over all levels of recursion.

Algorithm: **DMATMINPERM**

Purpose: Determine a mapping from the columns to the rows of the input matrix which minimizes the distance function represented by the input matrix

Input: **GLOBDMIN** The global minimum value of the distance; only mappings with distance less than this value will be returned.

DSOFAR The total accumulated distance by the algorithm from mappings generated from previous levels of recursion

DMAT A matrix of pairwise distances relating rows and columns

Output: **PERM** A 1-1 mapping from the columns into the rows of **DMAT**

DMIN The value of the distance associated with returned mapping

Initialization: **GLOBDMIN** $\leftarrow \infty$.

DSOFAR $\leftarrow 0$.

DMAT $\leftarrow$ d.

```
Subroutine: [PERM,DMIN] ← DMATMINPERM ( DMAT,DSOFAR)
if only one column in DMAT,
  begin
      PERM ← row corresponding to minimum value in DMAT;
      DMIN ← minimum value in DMAT;
      if DMIN+ DSOFAR < (GLOBDMIN)
        then (GLOBDMIN) ← DMIN+ DSOFAR;
      exit
  end
else
  begin
      MINEL ← value of the minimum of the first column of DMAT;
      CANDROW ← row corresponding to the minimum value of the first col-
                umn of DMAT;
      if MINEL + DSOFAR < GLOBDMIN,
        begin
            [CANDPERM, CANDDMIN ] ← DMATMINPERM (
                DMAT minus first column and row # CAN-
                DROW, DSOFAR + MINEL);
          PERM ← CANDROW concatenated onto CANDPERM;
          DMIN ← CANDDMIN + MINEL ;
          while there are rows left to consider in the first column of
              DMAT and MINEL + DSOFAR < GLOBDMIN,
              begin
                  [CANDPERM, CANDDMIN ] ← DMATMIN-
                      PERM ( DMAT minus first col-
                      umn and row # CANDROW,
                      DSOFAR + MINEL);
                if DMIN > CANDDMIN + MINEL,
                begin
                    PERM ← CANDROW concatenated
                        onto CANDPERM;
                    DMIN ← CANDDMIN + MINEL ;
                end
                Eliminate row CANDROW from consideration;
                MINEL ← value of the minimum from the remaining
                        rows in the first column of DMAT;
```

CANDROW ← row corresponding to the minimum value of the remaining rows in the first column of **DMAT**;

end

end

end

Notice that the basic recursive step consists of choosing the most promising candidate from Λ_B for matching to the element of Λ_A which corresponds to the first column and passing a reduced d-matrix to a lower level of recursion. The choice of the candidate from Λ_B is based on its distance to the element of Λ_A corresponding to the first column. The reduced matrix consists of the input matrix with the row and column which correspond to the chosen incremental match stripped away. In this way matrix reduction is equivalent to incremental formation of a candidate match.

At each step in the recursion, the algorithm effectively orders the elements of Λ_B by considering them in increasing magnitude of the distance to the element of Λ_A corresponding to the first column. By comparing the sum of the distance associated with the incremental candidate match and the total accumulated distance (**MINEL** + **DSOFAR**) to the global minimum distance (**GLOBDMIN**) it can be determined if there exists the possibility of achieving a lower overall distance. By rejecting all matches derived from such a hopeless candidate match, a significant reduction in the number of matches which are considered can be made. Furthermore, by ordering the elements of Λ_B at each level of recursion, this pruning ability is enhanced.

Note that elements of the first M_b rows of the d-matrix consist of one term from the first sum, one term from the second sum, and one term from the fifth sum of Equation (A.2). Furthermore, elements from the remaining rows of the d-matrix are strictly from the first sum. Therefore, by choosing different candidate incremental matches, one is implicitly choosing a different set of terms from the first, second and fifth sum in Equation (A.2). Furthermore, by considering only 1-1 matches (or possibly 1-$\emptyset$), no terms from the third or fourth sum can be included in the overall match distance.

A.2 The Generalized d-matrix Algorithm

The matches formed by the generalized d-matrix algorithm are potentially complex. The generalized d-matrix algorithm can form matches between two parametric decompositions in which any number of modes from one of the parametric decompositions can be matched to any number of modes in the other parametric decomposition. The only restriction regarding the formation of a match is that a mode may appear in only one match between subsets of modes of the two parametric descriptions.

In order to form such general matches in an optimal way, not only must there exist information regarding the pairwise distance between the modes of the two parametric decompositions but also information regarding the distance between an arbitrary subset of modes of one parametric decomposition and an arbitrary subset of modes of the other parametric decomposition. This information is available in the form of inner products

between the various modes. These inner products are related to the "cross-terms" in Equation (A.2). In addition the generalized d-matrix algorithm calculates the optimal value of the relative range shift between the two parametric decompositions. Therefore, this algorithm is not only applicable to estimation of the statistics of a parametric decomposition modes but also to the recognition of radar targets.

A matching function, $\mathcal{M}$, as is used to describe the non-symmetric matches generated by the d-matrix algorithm is, in the current case, not a 1-1 relationship between the elements of the two parametric decompositions. The type of matches which are generated by the generalized d-matrix algorithm are better described in terms of a partition (see Chapter 3) of the union of the parametric decomposition Λ_A with the parametric decomposition $(-\Lambda_B)$ (i.e. the modes of Λ_B have been replaced with their negatives.) Let $\mathcal{P}$ be a partition of $\Lambda_A \cup (-\Lambda_B)$ $(\mathcal{P} = \{\pi \mid \pi \subseteq \Lambda_A \cup (-\Lambda_B), \pi \neq \emptyset\} \cup \mathcal{P} = \Lambda_A \cup (-\Lambda_B))$. Each element of a partition implies a "match" between elements of the Λ_A and Λ_B. For example, supposing $\vec{v}_{a_2}, \vec{v}_{a_3} \in \Lambda_A$ and $-\vec{v}_{b_2} \in -\Lambda_B$ and a particular element of a partition is $\pi = \{\vec{v}_{a_2}, \vec{v}_{a_3}, -\vec{v}_{b_2}\}$ then this element implies a match between the combination of elements $\vec{v}_{a_2}$ and $\vec{v}_{a_3}$ in Λ_A to $\vec{v}_{b_2}$ in Λ_B. The residual energy associated with this match is then $|\vec{v}_{a_2} + \vec{v}_{a_3} - \vec{v}_{b_2}|^2$. The quantity which is to be minimized is therefore

$$\sum_{\pi \in \mathcal{P}} \left| \sum_{\vec{v}_a \in \pi \cap \Lambda_A} \vec{v}_a + \sum_{-\vec{v}_b \in \pi \cap (-\Lambda_B)} -\vec{v}_b \right|^2 . \qquad (A.7)$$

over all partitions of $\Lambda_A \cup (-\Lambda_B)$. Elements of Λ_B are replaced with their negatives so that the distinction between elements of Λ_A and Λ_B can be ignored. Minimization of the distance is now in terms of "combining" elements of the composite parametric decomposition.

Note that the algorithm is still based on the concept of minimization of residual energy. The difference lies in the more general way in which modes of one parametric decomposition may be used in "groups" to cancel the energy in "groups" of modes from the other parametric decomposition and in the fact that cancellation of energy in the modes of Λ_B is as important as cancellation of energy in the modes of Λ_A.

The use of the more general sense of matching allows the matcher to account for circumstances in which two noiseless modes are estimated as one mode in a noise realization. The reverse case (when one mode splits into two modes due to noise perturbation) may also be accounted for with this matching algorithm. In addition, since the distance function accounts for nil-mapping of elements of both parametric decompositions, complications due to non-symmetric matching are avoided.

A.2.1 Operation of the Generalized d-matrix Algorithm

Carrying out the minimization of (A.7) is the objective of the generalized d-matrix algorithm. This is accomplished by searching over all partitions of the set $\Lambda_A \cup (-\Lambda_B)$. Partitions of $\Lambda_A \cup (-\Lambda_B)$ determine and are determined by the current value of the operational matrix.

Define the partition $\mathcal{P}_0$ as $\{(\vec{v}_{a_1}), (\vec{v}_{a_2}), \ldots, (\vec{v}_{a_{M_a}}), (-\vec{v}_{b_1}), (-\vec{v}_{b_2}), \ldots, (-\vec{v}_{b_{M_b}})\}$. This partition is referred to as the "total-nil-map" partition since the matches implied by this partition are all "nil-maps", i.e. each individual element of Λ_A and Λ_B are mapped to "nil" (or $\emptyset$).

The algorithm begins by calculating the generalized d-matrix associated with the total-nil-map partition. For an arbitrary partition of $\Lambda_A \cup (-\Lambda_B)$, $\mathcal{P} = \{\pi_1, \pi_2, \ldots, \pi_M\}$, the associated generalized d-matrix can be computed by first forming the matrix $\boldsymbol{\mathcal{P}}$. This matrix consists of a number of *composite* modes, one for each element of the partition $\mathcal{P}$. Recall that an element of $\mathcal{P}$ is a set of vectors, each vector corresponding to a mode in Λ_A or Λ_B. For an arbitrary element, $\pi \in \mathcal{P}$ the corresponding composite mode, $\vec{\pi}$, is defined as the vector sum of the modes contained in π

$$\vec{\pi} = \sum_{\vec{v} \in \pi} \vec{v}. \tag{A.8}$$

Therefore if we define $\vec{\pi}_i$ as the composite mode corresponding to the partition element $\pi_i \in \mathcal{P}$, then the matrix $\boldsymbol{\mathcal{P}}$ is defined as

$$\boldsymbol{\mathcal{P}} = [\vec{\pi}_1, \vec{\pi}_2, \ldots, \vec{\pi}_M]. \tag{A.9}$$

The generalized d-matrix corresponding to the partition $\mathcal{P}$, denoted $\delta_{\mathcal{P}}$, is defined as

$$\delta_{\mathcal{P}} = \boldsymbol{\mathcal{P}}^H \boldsymbol{\mathcal{P}}. \tag{A.10}$$

The generalized d-matrix is Hermitian since $\delta_{\mathcal{P}}^H = (\mathcal{P}^H\mathcal{P})^H = \mathcal{P}^H\mathcal{P} = \delta_{\mathcal{P}}$. Furthermore, since there are M columns in $\mathcal{P}$ the matrix δ is of dimension $M \times M$. The element of the generalized d-matrix in row i column j is therefore the inner product of the composite mode $\vec{\pi}_i$ and $\vec{\pi}_j$, or

$$\delta_{\mathcal{P}}(i,j) = \vec{\pi}_i^H \vec{\pi}_j. \tag{A.11}$$

Recall that the total-nil-map partition includes all the of $\Lambda_A \cup (-\Lambda_B)$, each in a separate partition element. Elements of the generalized d-matrix associated with this partition, $\delta_{\mathcal{P}_0}(i,j)$ are given as

$$\delta_{\mathcal{P}_0}(i,j) = \begin{cases} \vec{v}_{a_i}^H \vec{v}_{a_j} & i,j \le M_a \\ -\vec{v}_{a_i}^H \vec{v}_{b_{(j-M_a)}} & i \le M_a,\ j > M_a \\ -\vec{v}_{b_{(i-M_a)}}^H \vec{v}_{a_j} & i > M_a,\ j \le M_a \\ \vec{v}_{b_{(i-M_a)}}^H \vec{v}_{b_{(j-M_a)}} & i > M_a,\ j > M_a \end{cases}. \tag{A.12}$$

where $\delta_{\mathcal{P}_0}(i,i)$ is the energy contained in the modes of $\Lambda_A \cup \Lambda_b$. If $i \le M_a$ then the mode corresponding to this row and column is from Λ_A and $\delta(i,i)$ is a term from the first sum of Equation (A.2). Likewise if $i > M_a$ then $\delta(i,i)$ is the energy from mode number $i - M_a$ from Λ_B and is therefore equivalent to a term from the second sum of Equation (A.2), *viz*

$$\delta_{\mathcal{P}_0}(i,i) = \begin{cases} |\vec{v}_{a_i}|^2 & i \le M_a \\ |\vec{v}_{b_{(i-M_a)}}|^2 & i > M_a \end{cases}. \tag{A.13}$$

Similarly, the off-diagonal terms, when summed with their symmetric counterparts, are terms from the third fourth or fifth sum of Equation (A.2). Therefore, assuming $i < j$

$$\delta_{\mathcal{P}_0}(i,j) + \delta_{\mathcal{P}_0}(j,i) = \delta_{\mathcal{P}_0}(i,j) + \delta_{\mathcal{P}_0}(i,j)^* \quad \text{(A.14)}$$

$$= 2\Re\{\delta_{\mathcal{P}_0}(i,j)\} \quad \text{(A.15)}$$

$$= \begin{cases} 2\Re\{\vec{v}_{a_i}^H \vec{v}_{a_j}\} & i < j \leq M_a \\ 2\Re\{\vec{v}_{b_{(i-M_a)}}^H \vec{v}_{b_{(j-M_a)}}\} & j > i > M_a \\ -2\Re\{\vec{v}_{a_i}^H \vec{v}_{b_{(j-M_a)}}\} & i \leq M_a \ \wedge \ j > M_a \end{cases} \quad \text{(A.16)}$$

The off-diagonal elements for the last case in Equation (A.16) (corresponding to the fifth sum in Equation (A.2)) are inner products between elements of the two parametric decompositions. As such they are referred to as being *inter-object*. Thus, these elements are therefore a function of the assumed range offset between the two objects. The other cases all represent intra-object inner products and as such are not a function of the assumed range offset. Elements of d-matrices generated under other partitions of $\Lambda_A \cup (-\Lambda_B)$ may not have the range dependent elements so nicely localized.

The generalized d-matrix algorithm is described under the assumption that the range offset between the two parametric decompositions is fixed. Thus, we initially assume that no elements of the generalized d-matrix are a function of the range offset. Issues relating to determination of the optimal range offset by the algorithm are addressed in Section A.2.2.

The initial value for the generalized d-matrix is to be input to the generalized d-matrix algorithm for formation of the correspondence between

the elements of the two parametric decompositions. This is done since a generalized d-matrix for an arbitrary partition can be generated from the initial value using a simple reduction procedure on the matrix $\delta_{\mathcal{P}_0}$. Before the generalized d-matrix algorithm is described, some properties of the generalized d-matrix are discussed. In particular, it is shown how the match quality function for the given partition is computed from the associated d-matrix. In addition a single step of the reduction procedure is described.

The value of the match quality function (A.7) can be easily computed from the generalized d-matrix. An arbitrary element of a given partition, π, is a subset of $\Lambda_A \cup (-\Lambda_B)$ therefore

$$\sum_{\vec{v}_a \in \pi \cap \Lambda_A} \vec{v}_a + \sum_{-\vec{v}_b \in \pi \cap (-\Lambda_B)} -\vec{v}_b = \sum_{\vec{v} \in \pi} \vec{v} \tag{A.17}$$

$$= \vec{\pi} \tag{A.18}$$

and the match quality function is given by

$$\sum_{\pi \in \mathcal{P}} \left| \sum_{\vec{v}_a \in \pi \cap \Lambda_A} \vec{v}_a + \sum_{-\vec{v}_b \in \pi \cap (-\Lambda_B)} -\vec{v}_b \right|^2 = \sum_{\pi \in \mathcal{P}} |\vec{\pi}|^2 \tag{A.19}$$

$$= \sum_{\pi \in \mathcal{P}} \vec{\pi}^H \vec{\pi} \tag{A.20}$$

$$= \sum_{i=1}^{M} \delta_{\mathcal{P}}(i,i) \tag{A.21}$$

$$= \mathrm{tr}[\delta_{\mathcal{P}}]. \tag{A.22}$$

An arbitrary partition of $\Lambda_A \cup (-\Lambda_B)$ can be formed from the initial partition, $\mathcal{P}_0$, by recursively replacing pairs of elements of the input partition with their union (in a selective fashion). It is for this reason, together

with the simplicity of the resulting procedure, that the operation matrix reduction step is defined by the union operation on the input partition.

Suppose that a given generalized d-matrix, $\delta_{\mathcal{P}}$ has associated with it the partition, $\mathcal{P} = \{\pi_1, \pi_2, \ldots, \pi_M\}$. Then the matrix of composite modes associated with the partition, $\mathcal{P}^{k,l} = \{\pi_1, \pi_2, \ldots, \pi_k \cup \pi_l, \ldots, \pi_{l-1}, \pi_{l+1}, \ldots, \pi_M\}$, in terms of the elements of $\boldsymbol{\mathcal{P}}$ is (note: the partition $\mathcal{P}^{k,l}$ has one fewer element than $\mathcal{P}$, i.e.$|\mathcal{P}^{k,l}| = M - 1$)

$$\boldsymbol{\mathcal{P}}^{k,l} = [\vec{\pi}_1, \vec{\pi}_2, \ldots, \vec{\pi}_k + \vec{\pi}_l, \ldots, \vec{\pi}_{l-1}, \vec{\pi}_{l+1}, \ldots, \vec{\pi}_M]. \qquad (A.23)$$

Thus, replacing the partition element π_k with the union of π_k and π_l in the partition $\mathcal{P}$ is equivalent to replacing the composite mode $\vec{\pi}_k$ by the sum of the composite modes $\vec{\pi}_k$ and $\vec{\pi}_l$ in the matrix $\boldsymbol{\mathcal{P}}$. If we define the $M \times M-1$ matrix $\Phi^{k,l}$ as the $M \times M$ identity matrix which has been column reduced as described above then

$$\Phi^{k,l}(i,j) = \begin{cases} 1 & i = j,\ i \leq l-1 \\ 1 & i = j+1,\ l+1 \leq i \leq M \\ 1 & i = l,\ j = k \\ 0 & \text{else} \end{cases} . \qquad (A.24)$$

Furthermore, post-multiplication by this matrix carries out column reduction on the matrix $\boldsymbol{\mathcal{P}}$ as described above so that

$$\boldsymbol{\mathcal{P}}^{k,l} = \boldsymbol{\mathcal{P}}\ \Phi^{k,l}. \qquad (A.25)$$

Likewise, pre-multiplying a matrix by $(\Phi^{k,l})^T$ replaces the k^{th} row by the sum of the k^{th} and l^{th} rows. The generalized d-matrix corresponding to $\mathcal{P}^{k,l}$ is therefore

$$\begin{aligned}
\delta_{\mathcal{P}}^{k,l} &= (\boldsymbol{\mathcal{P}}^{k,l})^H(\boldsymbol{\mathcal{P}}^{k,l}) && \text{(A.26)}\\
&= (\boldsymbol{\mathcal{P}}\ \Phi^{k,l})^H(\boldsymbol{\mathcal{P}}\ \Phi^{k,l}) && \text{(A.27)}\\
&= (\Phi^{k,l})^T(\boldsymbol{\mathcal{P}})^H(\boldsymbol{\mathcal{P}})\Phi^{k,l} && \text{(A.28)}\\
&= (\Phi^{k,l})^T\delta_{\mathcal{P}}\Phi^{k,l}. && \text{(A.29)}
\end{aligned}$$

The generalized d-matrix, $\delta_{\mathcal{P}}^{k,l}$, can therefore be derived from $\delta_{\mathcal{P}}$ with the following procedure:

1. Replace the k^{th} row of $\delta_{\mathcal{P}}$ by the sum of the k^{th} row and the l^{th} row.
2. Replace the k^{th} column of this resulting matrix by the sum of the k^{th} column and the l^{th} column.

This matrix reduction procedure forms the basis for the recursive step in the generalized d-matrix algorithm. In terms of the implied partition, a single step of generalized d-matrix reduction is equivalent to forming the union of two elements of the implied input partition. Formation of a partition is placed in terms of sequential "unioning" since replacing two partition elements with their union is an operation which preserves the partition property. Furthermore, any match (partition) can be derived from recursively unioning elements of the initial partition, $(\mathcal{P}_0)$. Therefore, a list of the unions (or "pairings") which are formed at each level of recursion

can serve as a representation of the implied partition. In this way, matrix reduction is equivalent to incremental match formation.

A further reason for using this method is that the increment in the match quality function associated with the reduced matrix is easily computed. The match quality function associated with the reduced matrix, $\delta_{\mathcal{P}}^{k,l}$, is

$$\begin{aligned} \mathrm{tr}[\delta_{\mathcal{P}}^{k,l}] &= |\vec{\pi}_1|^2 + |\vec{\pi}_2|^2 + \cdots + |\vec{\pi}_k + \vec{\pi}_l|^2 + \cdots + |\vec{\pi}_{l-1}|^2 + \\ &\quad |\vec{\pi}_{l+1}|^2 + \cdots + |\vec{\pi}_M|^2 \qquad \text{(A.30)} \\ &= 2\Re\{\vec{\pi}_k^H \vec{\pi}_l\} + \sum_{i=1}^{M} |\vec{\pi}_i|^2 \qquad \text{(A.31)} \\ &= 2\Re\{\vec{\pi}_k^H \vec{\pi}_l\} + \mathrm{tr}[\delta_{\mathcal{P}}] \qquad \text{(A.32)} \\ &= 2\Re\{\delta_{\mathcal{P}}(k,l)\} + \mathrm{tr}[\delta_{\mathcal{P}}]. \qquad \text{(A.33)} \end{aligned}$$

Therefore, twice the real part of the off-diagonal elements of the generalized d-matrix represents the incremental change in the match quality function which occurs if the partition elements which correspond to this element are chosen for union formation. While deciding on the element of the operational matrix which optimizes the incremental change in the match quality function is a strategy which is "1-step" optimal, a sequence of k such decisions are not necessarily k-step optimal (i.e. they do not necessarily optimize the match quality function after k recursions). Thus the generalized d-matrix algorithm must, in general, be recursive.

As was the d-matrix algorithm, the generalized d-matrix algorithm is also described in terms of a recursive subroutine. Input to the algorithm is the initial value of the generalized d-matrix, $\delta_{\mathcal{P}_0}$ for the variable **GDMAT**.

The output is the optimizing match, **LOP**, this variable is in the form of a list of unions of the elements of the recursively formed partitions and the associated value of the match quality function.

Algorithm: **GDMATMINPART**

Purpose: Determine a partition of the elements of the input set which minimizes the distance function represented by the input matrix

Input: **GDMAT** A matrix of inner products among the elements which correspond to the rows and columns

INLOP A list of pairings which have been applied to the initial generalized d-matrix, so far

Output: **LOP** A list of pairings of the elements of the recursively formed partitions. Used to represent the optimal partition.

DMIN The value of the distance associated with returned partition implied by **LOP**

Initialization: **GLOBDMIN** $\leftarrow \infty$.

GDMAT $\leftarrow \delta_{\mathcal{P}_0}$.

Subroutine: [**LOP**,**DMIN**] $\leftarrow$ **GDMATMINPART** (**GDMAT**, **INLOP**)

if only one element in **GDMAT**,
  begin
    LOP $\leftarrow$ **INLOP**;
    DMIN $\leftarrow$ **GDMAT**;
  end
else
  begin
    DMIN $\leftarrow$ tr[**GDMAT**];
    NPARTELS $\leftarrow$ Number of rows or columns in **GDMAT** (equivalent to number of implied partition elements);
    if **INLOP** is empty,
      begin
        i= 1;
        k= 2;
      end

```
else
  begin
        (i,k)= last pairing in INLOP;
        if k > NPARTELS
          begin
                i ← i+ 1;
                k ← i+ 1;
          end
  end
while i ≤ NPARTELS-1,
      begin
            while k ≤ NPARTELS,
                  begin
                        REDGDMAT ← (Φ^{i,k})^T GDMAT Φ^{i,k} ;
                        OUTLOP ← the pair (i, k) appended to INLOP;
                        [POSSLOP, POSSDMIN] ← GDMATMIN-
                                    PART ( REDGDMAT, OUT-
                                    LOP );
                        if POSSDMIN < DMIN
                          begin
                                LOP ← POSSLOP;
                                DMIN ← POSSDMIN;
                          end
                        k ← k+ 1;
                  end
            i ← i+ 1;
            k ← i+ 1;
      end
end
```

A.2.2 Properties of the Generalized d-matrix algorithm

The algorithm considers all possible pairings of elements of the implied partition in such a way that partitions are uniquely enumerated. At a given level of recursion, the algorithm considers elements of the upper triangular portion of the input generalized d-matrix (and thus, implicitly, pairings of elements of the implied partition) in "row-major" [40] order. Elimination of duplication of partitions is accomplished by commencing the search over pairings of the input partition (i.e. elements of the d-matrix) at the pairing which was made in the previous level of recursion. This is the purpose of the statements regarding the index variables **i** and **k**.

The topic of "range alignment", that is, determination of the value of the physical offset between the two targets has been delayed until this point to allow for clarity of presentation. The assumption that has been made to this point is that the relative position of the two targets is fixed at some value. This, of course, is not a valid assumption for the target identification problem since during operation of an actual radar system, a given target can occur at any location within the unambiguous range cell.

If we assume that the relative location of the target represented by parametric decomposition Λ_A is unknown with respect to the target represented by parametric decomposition Λ_B then several modifications must be made to the generalized d-matrix algorithm. In particular, elements of the initial value of the generalized d-matrix which correspond to inter-target pairings

become dependent upon the assumed range offset. The adaptation made to the generalized d-matrix algorithm to account for this is to evaluate each of these inter-target elements of the operational matrix at the optimizing value of range offset (i.e. that range offset at which twice the real part of these elements are minimum). Thus the off-diagonal elements of the generalized d-matrix represent an upper bound on the amount by which the match quality function decreases if the union of partition elements to which they correspond is chosen.

Furthermore, as operational matrices are formed from the initial generalized d-matrix, the inner products between two range offset dependent elements are summed with other range offset dependent elements. As this occurs, the values of these "summed" inner products must be optimized with respect to range offset. In this way, the value of the off-diagonal elements of the operational matrices always represent optimizing values with respect to match quality function reduction. Therefore, schemes by which the total number of considered partitions is reduced through examination of off-diagonal elements (pruning of the search space) remain valid.

Recall that the match quality function is the sum of the on-diagonal terms of the operational matrix. Since this function represents the residual energy associated with a match between elements of the two parametric decompositions, the range offset between the two targets must apply to all terms of the distance function associated with the given matches between the elements of the two parametric decompositions. Therefore, when range

offset dependent elements have propagated into on-diagonal elements of the operational matrix, then match quality function optimization (with respect to the range offset) must be accomplished "simultaneously" with other on-diagonal range offset dependent elements under a single optimizing value for range offset.

The inclusion of range offset dependencies implies that matrix reduction cannot be described by a simple sums of columns and rows of the operational matrix. Instead an optimization step must occur whenever two range offset dependent elements are joined. This additional computation brought by this complication can be partially offset, however, by comparison of the optimizing values of the range offset to reduce the number of partitions which are evaluated. This is what is referred to as relational consistency in the text of this document.

Recall that the range offset dependent elements of operational matrices are evaluated at an optimizing value for the range offset. In addition, the on-diagonal elements of an operational matrix are evaluated at a common value of the range offset parameter which optimizes the match quality function. The value for the range offset parameter represents the physical offset between the two targets. For the on-diagonal elements this value is optimized under the match between the elements of the two parametric decompositions implied by the input partition. For the off-diagonal elements of the operational matrices the values of the range offsets for the represent the physical range offset between the two targets *if* the match implied by the

given off-diagonal were to contain all the range dependent terms. Thus, a comparison of the value of the range offset associated with the on-diagonal terms with the range offset associated with a given off-diagonal element can determine if the increment in the match implied by the off-diagonal element of the operational matrix is consistent with the match implied by the on-diagonal elements of the input operational matrix. This comparison of range offsets is defined in terms of the width of the sum of the on-diagonal elements as a function of the range offset as described in Chapter 4.

While the algorithms have been given in terms of the vector form of the modes of the parametric decompositions, there exist closed form expressions for the elements of all the matrices in terms of the complex scalars which make up the parametric decompositions. The vector forms of the modes were used for clarity of presentation.

A.3 Summary

The d-matrix algorithm minimizes the pairwise distance between elements of the two parametric decompositions. This algorithm implicitly accounts for all energy in Λ_A while there is no "penalty" for not using elements of Λ_B. This algorithm could be implemented with range alignment, as is done with the generalized d-matrix algorithm. The basic recursive step in the d-matrix algorithm works is an elimination of a column of the operational matrix.

The generalized d-matrix, on the other hand works by "incorporating" a row and column of the operational matrix into a reduced order matrix. The match quality function associated with this algorithm accounts for all energy in both parametric decompositions, i.e. it "penalizes" for not making assignments to elements of either decomposition. In this way, the generalized d-matrix accomplishes a symmetric match.

Bibliography

[1] B. Bhanu, "Automatic target recognition: State of the art survey," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-22, No. 4, pp. 364-379, July 1986.

[2] E.K. Walton, F.D. Garber, D.L. Moffatt, A.J. Kamis, N.F. Chamberlain, and O. Snorrason, "Radar target classification studies," Final Report 717220-1, The Ohio State University ElectroScience Laboratory, November 1986.

[3] J.S. Chen and E.K. Walton, "Comparison of two target classification techniques," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-22, No. 1, pp. 15-22, 1986.

[4] A.A. Ksienski, Y.T. Lin and L.J. White, "Low-frequency approach to target identification," *Proceedings of the IEEE*, Vol. 63, No. 12, pp. 1651-1660, December 1975.

[5] G.S. Sebestyen, *Decision Making Processes in Pattern Recognition*, Macmillan, New York, 1962.

[6] H.L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I*, Wiley, New York, 1968.

[7] K.S. Fu, *Syntactic Pattern Recognition and Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

[8] H.G. Barrow and R.J. Popplestone, "Relational Descriptions in Picture Processing," in *Machine Intelligence 6*, B. Meltzer and D. Michie, Editors, pp. 377-396, University Press, Edinburgh, 1971.

[9] R. Wilson, "Is vision a pattern recognition problem," in *Pattern Recognition: 4th International Conference*, J. Kittler, Editor, pp. 1-25, Springer, 1988.

[10] V.V. Mottl' and I.B. Muchnik, "Linguistic analysis of experimental curves," *Proceedings of the IEEE*, Vol. 67, No. 5, pp. 714-736, May 1979.

[11] O.S. Sands and F.D. Garber, "Pattern representations and syntactic classification of radar measurements of commercial aircraft," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-12, No. 2, pp. 204-211, February 1990.

[12] S.W. Shah and R.J.P. deFigueiredo, "Structural processing of waveforms as trees," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-38, No. 2, pp. 328-339, February 1990.

[13] R.S. Ledley, L.S. Rotolo, T.J. Golab, J.D. Jacobsen, M.D. Ginsberg and J.B. Wilson, "Fidac: Film Input to Digital Automatic Computer and Associated Syntax-Directed Pattern-Recognition Programming System," in *Optical and Electro-Optical Information Processing*, J.T. Tippet, D. Beckowitz, L. Clapp, C. Koester, and J.A. Vanderburgh, Editors, pp. 591-613, MIT Press, Cambridge, MA, 1965.

[14] A.V. Aho and J.D. Ullman, *The Theory of Parsing, Translation and Compiling, Vol. 1*, Prentice-Hall, Englewood Cliffs, NJ, 1972.

[15] R.M. Haralick and L.G. Shapiro, "The consistent labeling problem, part i," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-1, No. 2, pp. 173-184, April 1979.

[16] L.E. Lipkin, W.C. Watt, and R.A. Kirsch, "The Analysis, Synthesis and Description of Biological Images," *Annals of the New York Academy of Science*, Vol. 128, pp. 984-1012, 1966.

[17] H.G. Barrow, A.P. Ambler, and R.M. Burstall, "Some techniques for recognizing structures in pictures," in *Frontiers of Pattern Recognition*, S. Watanabe, Editor, pp. 1-30, Academic Press, New York, 1972 (From a paper delivered at the International Conference on "The Frontiers of Pattern Recognition," Honolulu, Hawaii, January 18-20, 1971).

[18] K.S. Fu, "On syntactic pattern recognition and stochastic languages," in *Frontiers of Pattern Recognition*, S. Watanabe, Editor, pp. 113-138, Academic Press, 1972 (From a paper delivered at the International Conference on "The Frontiers of Pattern Recognition," Honolulu, Hawaii, January 18-20, 1971).

[19] L.G. Shapiro and R.M. Haralick, "Structural Descriptions and Inexact Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 5, pp. 504-519, September 1981.

[20] M.A. Eshera and K.S. Fu, "A graph distance measure for image analysis," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-14, No. 3, pp. 398-408, June 1984.

[21] K.S. Fu, "A step towards unification of syntactic and statistical pattern recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 2, pp. 200-205, March 1983.

[22] W.H. Tsai and K.S. Fu, "Error-correcting isomorphisms of attributed relational graphs for pattern analysis," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-9, No. 12, pp. 757-768, December 1979.

[23] W.H. Tsai and K.S. Fu, "Attributed grammar - A tool for combining syntactic and statistical approaches to pattern recognition," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-10, No. 12, pp. 873-885, December 1980.

[24] K.L. Boyer and A.C. Kak, "Structural stereopsis for 3-D vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-10, No. 2, pp. 144-166, March 1988.

[25] A. Sanfeliu and K.S. Fu, "A distance measure between attributed relational graphs for pattern recognition," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol SMC-13, No. 3, pp. 353-362, May 1983.

[26] A.K.C. Wong and M. You, "Entropy and distance of random graphs with application to structural pattern recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-7, No. 5, pp. 599-609, September 1985.

[27] A.K. Wong and D.E. Ghahraman, "Random graphs: Structural-contextual dichotomy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 4, pp. 341-348, July 1980.

[28] G.Y. Tang and T.S. Huang, "A syntactic-semantic approach to image understanding and creation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-1, No. 2, pp.135-144, April 1979.

[29] L.G. Shapiro, "Structural descriptions and inexact matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 5, pp. 504-519, September 1981.

[30] C.W. Chuang and D.L. Moffatt, "Natural resonances of radar targets via Prony's method and target discrimination," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-12, No. 5, pp. 583-589, September 1976.

[31] M.I. Skolnik, *Introduction to Radar Systems*, McGraw-Hill, New York, 1980.

[32] E.K. Walton and J.D. Young, "The Ohio State University compact radar cross-section measurement range," *IEEE Transactions on Antennas and Propagation*, Vol. AP-32, No. 11, pp. 1218-1223, November 1984.

[33] R. Carriere and R. Moses, "High resolution radar target modeling using a modified Prony estimator," Technical Report 718048-11, The Ohio State University ElectroScience Laboratory, April 1989.

[34] L. Marple, *Digital Spectral Analysis with Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

[35] R. Kumaresan and D.W. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-30, No. 6, pp. 833-840, December 1982.

[36] M.A. Rahman and K. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-35, No. 10, pp. 1440-1454, October 1987.

[37] E.M. Kennaugh and D.L. Moffatt, "Transient and impulse response approximation," *Proceedings of the IEEE*, Vol. 53, No. 8, pp. 893-901, August 1965.

[38] C.F. Gerald, *Applied Numerical Analysis*, Addison Wesley, Reading, MA, 1978.

[39] D.H. Ballard and C.M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

[40] M.J. Augenstein and A.M. Tenebaum, *Data Structures and PL/I Programming*, Prentice-Hall, Englewood Cliffs, NJ, 1979.